Influencing factors of nasal flapping in English

Luca Garai

ELTE BTK Nyelvtudományi Doktori Iskola

garailuc@gmail.com

Abstract: The study is concerned with the flapping of intervocalic /nt/ and aimed to determine whether lexical frequency plays a significant role in flap production, whether male speakers are more likely to produce flaps than female speakers, and whether speech rate influences the likelihood of flap production. A set of samples was collected with words containing intervocalic /nt/ with the same preceding and following vowel, and then annotated in Praat to gather segment durations. The results were evaluated using descriptive statistics and logistic regression. The data show that higher lexical frequency positively influences the likelihood of flap production, that male speakers produce flaps more frequently than females, and that shorter word duration is associated with a higher likelihood of flapping. The results shed light on the specific nature of nasal flapping, as well as confirm some previous theoretical notions regarding the influencing factors of alveolar lenition.

https://doi.org/10.18135/Alknyelvdok.2021.15.7
In: Gráczi Tekla Etelka – Ludányi Zsófia (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2021.* Budapest: Nyelvtudományi Kutatóközpont. 2021. (Sorozatszerkesztő: Váradi Tamás.)
ISBN 978-963-9074-94-1.

1 Introduction

The process of flapping in English has been a heavily discussed topic in linguistics for almost a century. According to de Jong (2011: 2711), the first records of what we call flapping today were made in the 1930s, when scholars described the appearance of differently articulated word-medial /t/s. However, as Drinka (1990: 113) points out, Elphinston had already observed the voicing of fortis alveolar plosives in 18th century "vulgar" London speech, citing examples such as "proddestant" and "pardner." In the first half of the 20th century, Kenyon (1994: 126) noted the presence of "voiced t" between other voiced segments in American English and observed that it must differ from /d/ due to the lack of confusion by American speakers between pairs such as putting–pudding.

Since then, numerous studies have been conducted regarding different aspects of flapping. Its neutralizing effect, its various possible environments and context, its behavior in other dialects of English, and the phonemes it affects are all among the most frequent topics of debate when it comes to flapping (Vaux, 2000). The current study mainly focuses on the target phonemes of flapping, specifically /n/ and nasal-plosive clusters /nt, nd/. Literature on the flapping of these segments is lacking, as most analyses are directed at the flapping of alveolar plosives. However, due to the evolution of the phenomenon since its appearance in the early 20th century, the need for empirical research has become substantiated regarding nasal flapping in order to broaden the current understanding of the flapping process.

1.1 Alveolar flapping in American English

There are numerous features regarding English flapping which have been concurrently observed by most scholars. One of them is the produced flap sound itself. Contrary to how Kenyon (1994) first

perceived the flap in the English language, it is not simply a voiced version of /t/. In fact, it is so different that "flap" is usually cited as a separate manner of articulation (Laver, 1994: 130). Analyses of the type of articulatory motion which is involved in the production of flaps have shown that high speed and brief contact between active and passive articulators characterize the process (Laver, 1994: 145). Steriade (2000: 323) agrees that the feature which distinguishes flapped allophones from fully articulated ones is their "extra short closure" duration. According to Zue and Laferrière (1979: 1044), the majority of flap sounds "vary in duration from 10 to 40 ms."

In this regard, some scholars differentiate between flaps and taps, on the basis of a different form of contact—sliding or ballistic. However, since there is not enough evidence to support a clear and profitable distinction between the two processes, the question of flaps versus taps is primarily terminological at this point (de Jong, 1998: 284). The present paper utilizes the terms "flap" and "flapping", simply because, even though there is some conflict between viewpoints, the majority of the literature uses the same phrases to refer to this phenomenon in the English language.

According to Klatt (1975: 699), voiced-voiceless plosive pairs exhibit different closure durations in similar intervocalic environments. Articulatory analyses have shown that flapped /d/s and /t/s do come close in terms of closure duration, but their assimilation is not complete (Kenyon 1994: 316–7; Zue and Laferrière 1979: 1042). Therefore, even though the audible difference becomes so negligible that it inhibits perceptual differentiation and listeners must decipher the underlying phoneme by using other cues, it can be said that flapping is not a completely neutralizing process.

There have been various acoustic analyses regarding the neutralizing effect of alveolar flapping in English, which investigate whether or not word pairs such as *bitter–bidder* become homophonous as a result of flapping. Two of the most often studied features of the

phenomenon in this regard are closure duration and vowel duration. Kiparsky (1982: 127) claims that flapping is a process of "contextual neutralization," meaning that it merges /t/ and /d/ only in the distinct environments it occurs in. Kenyon (1994: 318) notes that /t/ flaps are shorter than /d/ flaps based on the comparison of spectrograms, thus the neutralization is only partial, occurring in the form of voicing. Furthermore, production studies analyzing the duration of the vowel preceding the flap have also shown distinct results: flapped /t/s are preceded by shorter vowels than flapped /d/s, which supports the claim that flapping yields incomplete neutralization (Braver, 2011: 37). Bermúdez-Otero (2004) attributes the differing vowel durations to Prefortis Clipping, which shortens the surface duration of vowels before voiceless plosives, including /t/. He claims that the reason why the vowel duration distinction is present even in flapped variants, where both /t/ and /d/ are voiced, is because Prefortis Clipping takes into account the underlying plosive phoneme, i.e. it is irrespective of the flapping process (5). Therefore, it can be said that flapping, at least on an acoustic level, does not neutralize word pairs such as *betting-bedding.*

One of the questionable features regarding alveolar flapping is the apparent variety between /t/ allophones ([ɾ], [t], and [ʔ]) even in previously established mandatory flapping environments. Steriade (2000) claims that discrepancies such as mandatory flapping in capitalistic but a fully articulated /t/ (with a voiceless closure phase and an explosive release) in militaristic are due to Paradigm Uniformity, meaning that the probability of flapping in these kinds of derived word forms corresponds to the pronunciation (and stress pattern) of their base forms, in the mentioned examples, to capital (/t/ between V́_V, flapped) and military (/t/ before secondary stress, unflapped), respectively. On the other hand, Patterson and Connine (2001: 267) attribute the intervocalic occurrence of all three allophones to lexical frequency and morphological complexity. They found that while 95% of high-frequency words exhibited flapping,

only 76% of low-frequency words did so, and that in monomorphemic words (e.g. city), 96% contained [ɾ] intervocalically, as compared to 63% in morphologically complex words, where the intervocalic /t/ was the result of an added suffix, such as parting (Patterson and Connine, 2001: 258–9). Therefore, they claim that the more a word with a possible flapping environment is uttered in everyday speech and the fewer morphemes it contains, the more likely speakers are to favor the production of [ɾ] above other /t/ allophones. Apart from the linguistic features of flapping, there are also some social connotations associated with it in English. Zue and Laferrière (1979: 1077) found that female speakers tend to produce much longer flaps than males and that women are also much more likely to articulate /nt/ fully, while men exhibit /n/-deletion almost exclusively, with either a flapped or an aspirated /t/. The authors attribute this difference to females preferring more careful, less reduced surface realizations than men. However, the gender distinction observed by Zue and Laferrière (1979) might be due to the formal context of their experiment, since their research did not focus on spontaneous utterances; instead, they instructed their subjects to read out a list of words in a controlled environment, which can lead to hypercorrection in most speakers. Nevertheless, potential gender differences in flap production should be examined.

1.2 Nasal flapping and post-nasal alveolar plosives

The presence of flapping after highly sonorous consonants has been documented since the late 1930s (Vaux, 2000). Kahn (1976: 64) notes that if /n/ precedes a flap, it cannot be consonantal, but instead only nasalizes the vowel before it and is then deleted. This way, in his analysis, the pronunciation of *winter* in casual speech is [wĩɾɚ], while careful speech preserves both the [n] and the [t]. Zue and Laferrière (1979: 1045) also find significant variation between the articulations of /t/ in post-nasal position. They mention fully ar-

ticulated (e.g. *minty* /'mɪntɪj/), deleted (e.g. *pentagon* /'pɛnə˩gʌn/), and flapped realizations as well, with the latter appearing in words where /n/ is deleted after nasalizing the preceding vowel, in words such as *antelope* /'æɾ̃˩ələwp/. However, Vaux (2000) counters this viewpoint by stating that most post-nasal alveolar plosives, instead of flapping, are deleted, thus yielding a nasal flap (formed from the remaining intervocalic nasal alveolar): [ɾ̃].

The flapping of /n/ is supported by Picard's (1997: 290) findings, who asserts that flapping "also applies to alveolar nasals in exactly the same environments as /t, d/" and that the nasal flap can appear as an allophone of /nt/ as well. He then compares the formation of homophonous *winter* and *winner* [wɪɾ̃ɚ], in order to propose a simpler flapping process than that of Jensen (1993). Instead of deleting /n/ first and then renasalizing the resulting flap, he opts for the deletion of the plosive and yielding the nasal flap directly from /n/ (compare fig. 1 and fig. 2).

LEXICAL REPRESENTATIONS	/wɪntər/	/wɪn+ər/
NASALIZATION	ĩ	ĩ
NASAL CONSONANT DELETION	∅	–
FLAPPING	ɾ	ɾ
OUTPUT	[wĩɾər]	[wĩɾər]
NASAL SPREADING	[wĩɾ̃ər]	[wĩɾ̃ər]

Figure 1: Jensen's derivation of *winter* and *winner* from Picard (1997: 291)

Picard's claims for a simpler analysis are convincing, however, in his assertion regarding the flapping of intervocalic nasals, he only refers to the fact that most previous analyses did not recognize /n/ as a possible subject to flapping (1997: 290). Furthermore, Vaux (2000) supports the flapping of /n/ in words like *twenty* with rule-ordering and referring to other theoretical claims, including those

of Picard (1997). The issue with these viewpoints advocating for nasal flapping lies in the (unjustified) premise that the intervocalic realization of the nasal alveolar (as in *winner*) can be phonetically classified as a flap (the properties of which are detailed above), as compared to a fully realized [n], and that this realization differs from flapped variants of /t, d/ .

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

Figure 2: Jensen's derivation of *winter* and *winner* from Picard (1997: 291)

Furthermore, whether *winner* and *winter* are, in fact, homophonous when the latter exhibits a flap between its two vowels is not backed up by Picard (1997) or Kahn (1976) with other than assumptions based on their own pronunciation. At the same time, while Zue and Laferrière (1979) provide an empirical basis for the realization of /nt/ clusters in an intervocalic environment and discover various possibilities, their research does not cover /n/ allophones and their conclusions follow the "/n/ deleted after nasalization, then /t/ flapped" path of flapped /nt/. Moreover, their study was conducted in the late 1970s and observed the pronunciation of only six subjects, rendering their results unsuitable for a 21st century analysis of the flapping phenomenon, which has been increasingly ubiquitous in the everyday use of English. The flapping of /nd/ clusters is not touched upon by any of the theoretical analyses dealing with nasal flapping. The observations made by Zue and Laferrière deem it impossible, as they note that "/n/ is almost never deleted" in this position, therefore the flap cannot be realized from the remaining plosive (1979: 1045). The present study takes into consideration the fact that the

deletion of the nasal may not be required for flapping, as well as the broadening set of elements which may undergo flapping, and it also analyzes samples with this cluster.

1.3 Previous findings and current hypotheses

The study conducted in the previous phase of the current research (Garai, 2019) shed light on various patterns regarding nasal flapping in American English. Its findings confirmed that the alveolar nasal /n/ has a distinct flapped allophone which is different from /t, d/ flaps in terms of segment durations, and that when flapped, the articulation of /nt/ clusters is more similar to that of /n/ flaps in terms of segment durations. One significant takeaway of the research regarding /nt/ flaps, however, is that the frequency of flapped realizations vis-à-vis fully articulated /nt/ clusters varies considerably between the words chosen for analysis, as can be seen in fig. 3.

In the case of *twenty* and *winter*, the overwhelming majority of samples were either flapped or unflapped (91% of tokens were flapped in *twenty* samples and 29% were flapped in the case of *winter*), whereas the ratio of flapped and unflapped counting samples was more balanced (57% flapped). Therefore, this consonant cluster is the ideal candidate for examining various hypotheses related to the phenomenon of alveolar flapping, as it is a currently emerging target of the lenition process in question. The current experiment seeks to evaluate the following three claims based on the analysis of samples with intervocalic /nt/ clusters in a similar phonetic environment:

(1) higher lexical frequency increases the likelihood of flap production;

(2) males are more likely to produce flaps than female speakers;

(3) faster speech increases the likelihood of flap production.

Hypothesis (1) is based on Patterson and Connine's (2001: 267) statements, while (2) reflects on the results of Zue and Laferrière's

(1979: 1047) research, as well as some observations made at the end of the previous study, and (3) aims to determine whether speech rate has an influence on flap frequency, since the sample set collected for analysis in the current study gives way to testing such a claim.

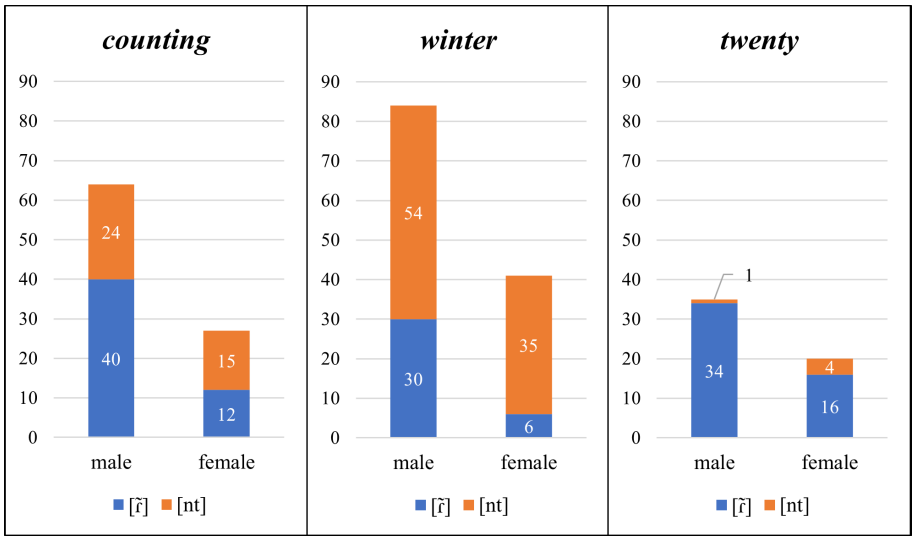

Figure 3: Distribution of /nt/ realizations across the three /nt/ words analyzed in the previous phase of the research, with regard to the gender of the speaker

2 Methods

As established above, the main objective of the current study was to analyze samples where /nt/ appears in a similar phonetic environment. Thus, the following four words were selected for analysis: *counting, mounting, accounting, amounting.* These words were chosen based on the findings in 1.3. Out of the three /nt/ words analyzed previously, counting showed the highest variability between

flapped and unflapped realizations, and as such, it was expected to supplement the evaluation of claims regarding various factors which may influence flap production. The other three words were selected because they provide the same phonetic environment for the nasal-plosive cluster in terms of adjacent vowels.

Fifty instances of each word were collected from the (semi-)spontaneous speech of various North American speakers in YouTube videos. `YouGlish.com` was utilized to search for each utterance. The samples were obtained using the output-recording capabilities of *Audacity* (Audacity Team, 2018). During sample collection, the main goal was to have samples with little to no interfering sound effects or other elements which may skew the data. Recordings with poor audio quality (due to the microphone used or the quality of the video itself) were also disqualified. It was also a priority to have a more balanced ratio between male and female subjects than in previous sample sets. The collection process did not allow for samples uttered by the same fifty subjects in all four sample sets. Although there was a small number of instances where multiple samples were taken from different utterances of the same speaker (this was possible for speakers who were easily identifiable and of whom there was a substantial amount of content uploaded to YouTube, e.g. Barack Obama). Flap production was not a factor when selecting potential samples, since the aim of the current research is to compare flapped and unflapped instances and their respective frequency. Therefore, the first fifty utterances which fit the aforementioned criteria were recorded, along with the gender of the speaker and, in the case of mounting and accounting, the syntactic category of the utterance.

Afterwards, all 200 samples were segmented and annotated in Praat (Boersma and Weenink, 2018). The durations extracted for analysis were closure duration (separately for /n/ and /t/ in unflapped instances, but these were later combined for comparison), word duration, and the duration of the homophonous /aʊntɪŋ/ portion of each sample. During the annotation process, the perceived

realization (fully articulated or flapped) of the /nt/ cluster in each sample was noted along with the durations measured, based on the author's individual judgment, as the original plan to conduct a perception experiment with various participants had to be postponed due to the pandemic. Samples where the two consonants could be distinguished from each other, both audibly and on the spectrogram, were marked as unflapped, including utterances where the plosive underwent voicing, and the cluster was realized as [nd] with a full occlusion before release. Meanwhile, samples with no distinct /n/ and /t/ realizations in the intervocalic closure were marked as flapped. An example pair of an unflapped and a flapped token of mounting are shown in spectral form in fig. 4, with the realization of the target /nt/ cluster indicated in the images.

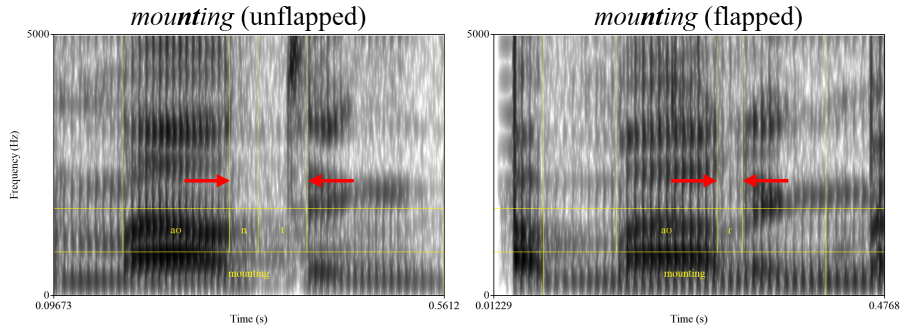

Figure 4: Spectrograms of the word mounting with unflapped /nt/ (left) and flapped /nt/ (right); arrows indicate the boundaries of the target consonant cluster.

The measurements were then exported into an Excel spreadsheet, and the statistical analysis took place in Python (Python Software Foundation, 2019), with the help of pandas (The pandas development team, 2021) and statsmodels (Seabold and Perktold, 2010). The mean, median, and standard deviation of closure durations in each sample set were calculated, separately for flapped and unflapped

instances; the confidence interval for the mean was 95% in all cases. Afterwards, logistic regression was used to determine the effect of the various factors on flap production. First, the relationship between consonant duration and perceived flap production was measured to determine the accuracy of the classification (flap status ∼ C duration). Then, a second logistic model was created to examine the effect of lexical frequency (based on *The Corpus of Contemporary American English, henceforth COCA* [Davies, 2008] and Google search results), the gender of the speaker (dummy coded), and speech rate (the duration of the homophonous portion of each sample) on flap production rate (flap status ∼ lexical frequency * gender * /aʊntɪŋ/ duration). Flap frequency in relation to syntactic category was also examined in the case of *mounting* and *accounting.*

3 Results

The data collected for analysis in the current study included samples with flapped and fully articulated realizations of /nt/ as well. Therefore, treating the gathered set of data as a single unit and conducting the statistical analysis on it as such would lead to the misrepresentation of values. It is expected that the measured durations of the intervocalic closure will fall into two (at least somewhat) distinguishable groups, since, as established by the literature (e.g. Steriade, 2000: 323) and previous findings, flapped realizations of alveolar consonants and nasal-plosive clusters exhibit much shorter closure durations than their fully articulated counterparts. Fig. 5 shows the distribution of the measured closure durations in all 200 samples. It can be seen on the histogram that one of the frequency peaks falls between the 10 to 40 ms range, which was established as the duration range for alveolar flaps by Zue and Laferrière (1979: 1044). The density of measured durations above 40 ms is less concentrated, but another frequency peak can be observed between 90 and 100 ms.

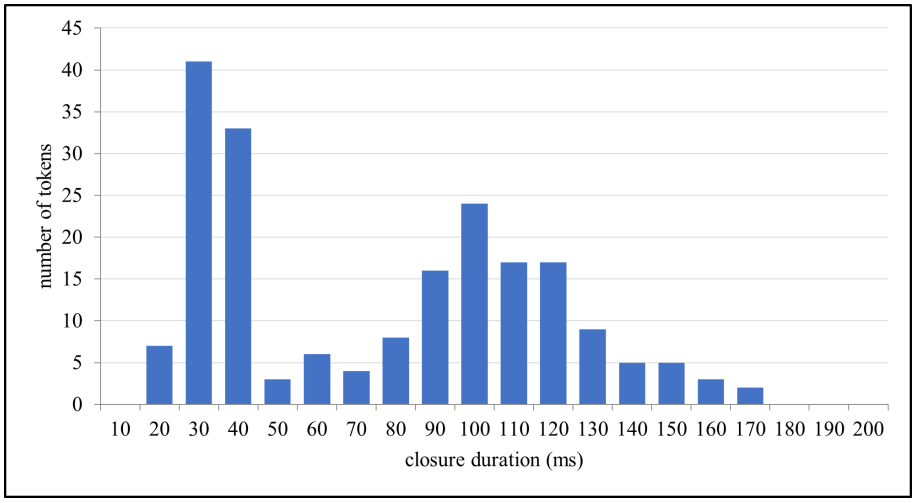

Figure 5: Histogram of closure durations across all /nt/ samples

During the annotation process, the perceived realization (unflapped or flapped) of /nt/ clusters in each sample was recorded, based on whether the author heard a fully articulated [nt] cluster with two distinct segments, or a single flap(like) sound. However, it would be unwise to categorize samples based solely on subjective judgment. Therefore, the effect of measured consonant duration on perceived flap production (denoted by 0 for unflapped samples and 1 for flapped ones) was measured using a logistic regression model. The results showed that the duration of the realized /nt/ cluster had a significant effect on whether the author perceived the realization as a flap or not ($z = 7.027; p < 0.001$). The mean duration of /nt/ clusters classified as flaps (29.51 ± 7.93 ms) was considerably lower than that of unflapped clusters (102.85 ± 24.7 ms). This signifies that if we take closure duration to be the primary distinguishing feature between flapped and fully articulated variants, flapped ones being substantially shorter (as argued by Steriade [2000: 323] among others), it can be assumed that the grouping of samples based on perceived flap production was accurate, and in general, the longer

the measured closure duration was in a given sample, the less likely it was to be perceived as a flap.

After arranging the samples into flapped and unflapped instances of each word, descriptive statistical analyses were conducted to calculate the mean and median of closure durations, the standard deviation, as well as the minimum and maximum value in each group of samples (word and flapping status). The calculated values, along with the number of samples in a given category, can be seen in table 1 below. The proportion of flapped realizations varies across the four words analyzed. While 27 out of 50 counting samples have a flapped /nt/ (amounting to 54% of the sample size) and 28 accounting samples are flapped (56%), the same can be said for only 13 samples in the mounting group (26%) and 16 amounting samples (32%). Overall, 84 of the 200 samples were perceived to contain flaps, which is 42% of all samples collected.

The mean and median values for flapped /nt/ closure durations each fall into the 10 to 40 ms range and differ considerably from those of unflapped instances. Similarly, the standard deviation of unflapped samples is substantially higher, allowing for a wider range for closure duration in fully articulated /nt/ realizations. Notably, the maximum value of flapped closure durations in all four sample groups is above the upper limit of the typical flap range, whereas the minimum value of unflapped closure durations among amounting samples is within the flap range. This, of course, does not mean that the duration range established in previous literature is necessarily incorrect, as there are bound to be some outliers in a set of data.

The visualization of the data presented in table 1 can be found in fig. 6 below. The previously established 40 ms upper limit for flap durations is marked with a red line, and it can be inferred that the majority of samples classified as flaps fall under this line, (only one flapped instance of each word had a consonant duration above 40 ms). The only overlap between flapped and unflapped realizations was observed in the case of amounting, but this was due to one

item in each group (a flapped cluster with a 54 ms duration and an unflapped /nt/ cluster which was 39 ms long), and the two sample groups (flapped and unflapped) are clearly distinguished by duration in the other three words.

Word	/nt/	**No. of tokens**	**Mean C duration (ms)**	**Median C duration (ms)**	**SD**	**Min. value**	**Max. value**
counting	[ɾ̃]	27	26.84	27.56	6.92	13.09	44.04
	[nt]	23	106.04	107.95	25.80	49.40	164.59
mounting	[ɾ̃]	13	28.11	23.29	10.57	20.57	59.01
	[nt]	37	111.30	107.27	24.08	75.44	164.90
accounting	[ɾ̃]	28	32.31	33.73	5.64	18.78	41.81
	[nt]	22	97.32	97.71	23.93	52.83	152.77
amounting	[ɾ̃]	16	30.27	28.99	9.38	16.84	53.74
	[nt]	34	95.06	93.84	22.69	38.54	154.07
all	[ɾ̃]	84	29.51	28.50	7.93	13.09	59.01
	[nt]	116	102.85	101.57	24.70	38.54	164.90

Table 1: Duration of intervocalic /nt/ across all four words sampled, categorized by perceived

3.1 Flap production based on lexical frequency

This subsection details the findings regarding the first one of these, which links lexical frequency and the likelihood of flap production based on previous claims made by Patterson and Connine (2001). For purposes of comparison, each of the four words was entered into the search function of The Corpus of Contemporary American English (Davies, 2008) and the number of results was recorded, along with the estimated Google search results of the same word. The collected data, along with the percentage of flapped realizations in

each word's sample set, can be seen in table 2. Out of the four words analyzed, accounting was found to be the most frequent both on the COCA website and in terms of Google search results. This corresponds to the fact that with 56%, it was the one with the highest percentage of flapped /nt/ clusters. The second most frequent word in COCA was counting, although it only placed third in terms of Google search results; this word had the second highest percentage of flapped realizations. The two words with lower flap production ratios, however, do not seem to follow the same principle. Even though *mounting* had over eight times as many search results in COCA than *amounting*, and over sixteen times as many Google search results, the percentage of flaps produced in *amounting* samples exceeded that of *mounting* by eight percent. The logistic model, which included the factors for all three hypotheses put forth (lexical frequency, gender of the speaker, /aʊntɪŋ/ duration), showed that in the case of lexical frequency, there was a significant effect on whether the given sample was articulated with a flap ($z = 5.441; p < 0.001$).

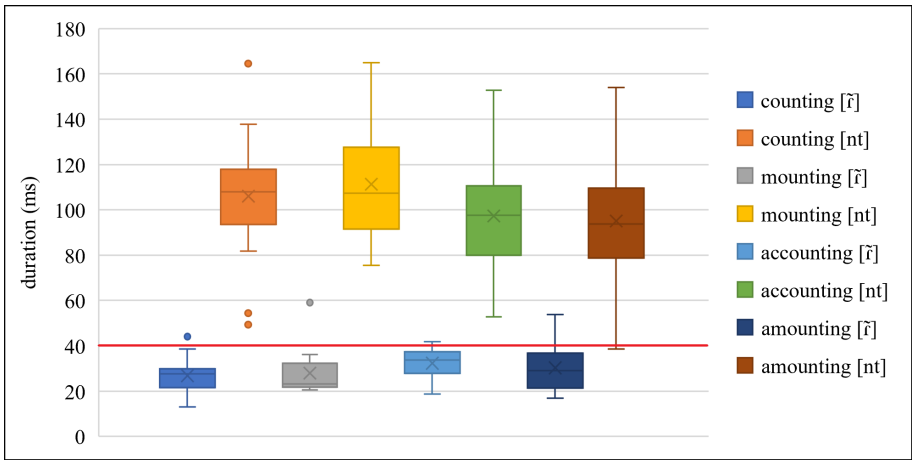

Figure 6: Comparison of closure durations in flapped and unflapped realizations of all four words analyzed

Word	% of flapped samples	COCA frequency	Google search frequency
counting	54%	18,610	265,000,000
mounting	26%	6,485	434,000,000
accounting	56%	19,396	501,000,000
amounting	32%	797	27,000,000

Table 2: Comparison of flap production ratio and the lexical frequency of each word

When collecting samples for *mounting* and *accounting*, the syntactic category of each sample was recorded. The utterances of *mounting* appeared in either a verbal (e.g. *mounting a horse*) or an adjectival (e.g. *mounting evidence*) position, while those of *accounting* were either verbs (e.g. *accounting for the loss*) or nouns (e.g. *I studied accounting*). Since COCA (Davies, 2008) has a feature to specify the part of speech of the occurrence during a search, the frequency of each category of the two words was gathered from the corpus. The displayed findings are visualized in fig. 7 and displayed in table 3 below. Based on the comparison between flapped versus unflapped occurrences and lexical frequency in each syntactic category of both words, it can be said that lexical frequency does play at least some role in determining the likelihood of flap production, since mounting samples were more likely to be flapped as adjectives than as verbs (albeit only slightly), and accounting samples were more often flapped as nouns than as verbs, both of which relations correspond to the respective lexical frequency of each category in COCA.

3.2 Flap production based on the gender of the speaker

In order to test the hypothesis of the current study which is concerned with the relationship between flap frequency and the gender of the speaker, we must assess the ratio of male to female speakers with

regard to flapped and fully articulated samples in the data. Overall, 65 of the 129 samples supplied by male speakers were flapped (50.39%), while 64 were unflapped (48.61%). As for female samples, only 19 out of 71 were flapped (26.76%) and 52 were unflapped (73.24%). The ratio of the two realizations among samples from male and female speakers for each of the four words can be seen in fig. 8.

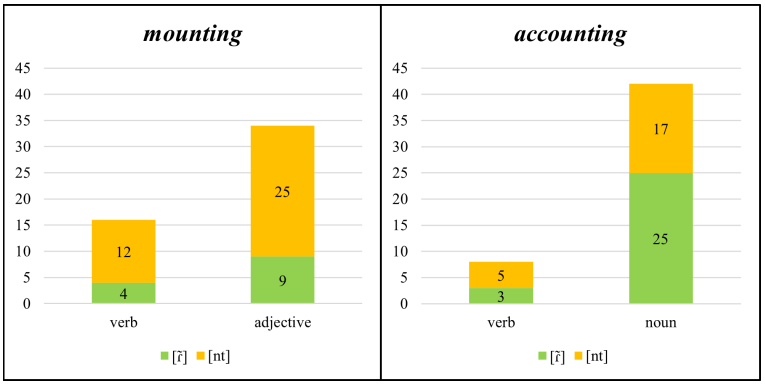

Figure 7: Distribution of flapped and unflapped mounting and accounting samples based on syntactic category

word	part of speech	flapped	unflapped	COCA frequency
mounting	verb	4 (25.0%)	12 (75.0%)	352
	adjective	9 (26.5%)	25 (73.5%)	1,686
accounting	verb	3 (37.5%)	5 (62.5%)	236
	noun	25 (59.5%)	17 (40.5%)	12,243

Table 3: Distribution of flapped and unflapped mounting and accounting samples based on syntactic category, with lexical frequency in COCA (Davies, 2008)

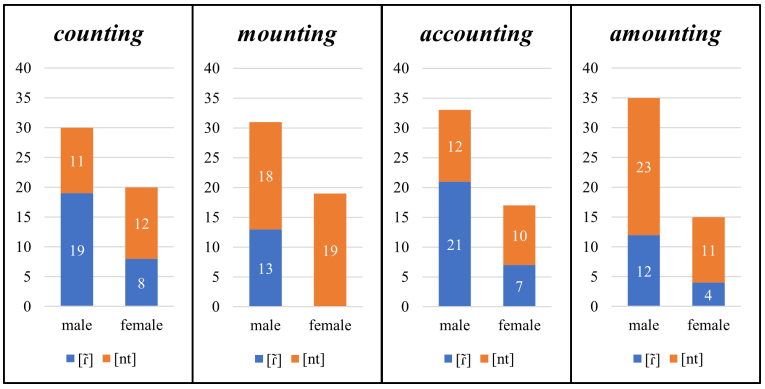

Figure 8: Distribution of flapped and unflapped realizations of /nt/ based on gender in all four sample sets

It can be deciphered from the data that male speakers are in an overall majority (making up 64.5% of all samples collected). However, as illustrated by the cumulative ratios, as well as in the separate sample groups above, female speakers were more likely to produce an unflapped articulation of /nt/ in all instances, outweighing males in the unflapped sample groups of counting and mounting. Furthermore, all utterances of mounting which were sampled from female subjects featured an unflapped realization of /nt/. The logistic model also showed a significant effect of speaker gender on the likelihood of flap production ($z = 2.414; p < 0.05$). This corresponds to the findings of Zue and Laferrière (1979), who concluded that female speakers are less likely to flap alveolar plosives intervocalically. Male speakers were more likely to produce flaps in counting and accounting over the fully articulated variant, but both male and female subjects had a tendency to not flap the /nt/ cluster in mounting and amounting. Overall, the findings of this subsection substantiate the claim that males are more likely to produce flapped /nt/ clusters intervocalically than female speakers; however, the sample size should be increased and more balanced in terms of gender to legitimately assess the hypothesis.

3.3 Flap production based on speech rate

The final claim hypothesized that speech rate has an influence on the flapping of /nt/ clusters, meaning that the faster a person speaks, the more likely they are to produce flaps. The way to test this claim under the circumstances of the current research is by looking at the relationship between word duration and the duration of the /nt/ cluster. The reason why the speech rate of longer utterances was not measured is because the heavy editing in many of the sampled YouTube videos would have skewed the results. The scatterplot in fig. 9 shows the relationship between word duration and /nt/ duration (the latter of which has been established as a considerable indicator of whether a segment is a flap or not) grouped by flapped and unflapped tokens.

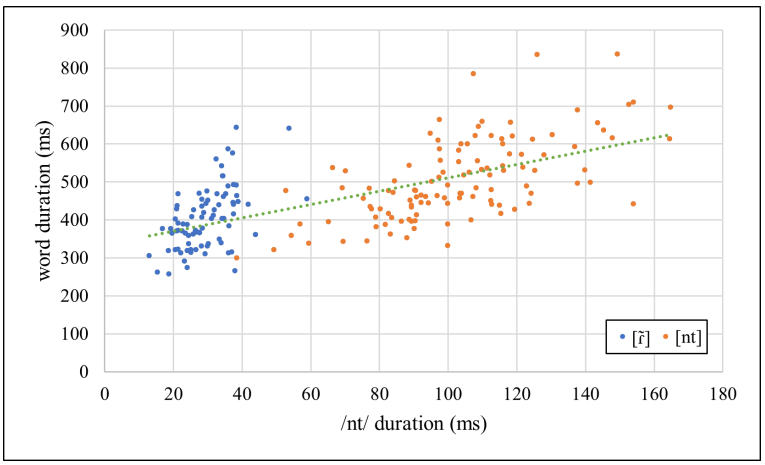

Figure 9: Scatterplot of the correlation between segment duration and word duration of flapped and unflapped /nt/ samples

It is important to note that the four words analyzed have varying syllable counts and/or begin with a differently articulated consonant. These differences may influence the data in a way which would lead

to inaccurate results. Therefore, the duration of segments before the vowel preceding /nt/ was subtracted from the word duration in each sample, resulting in the duration of the /aʊntɪŋ/ portion of all 200 samples, creating underlyingly homophonous tokens. This /aʊntɪŋ/ duration was used in the logistic regression model. The results show that the duration of this part of the word had a significant effect on flap production likelihood ($z = 5.223; p < 0.001$), and based on fig. 10 below, we can infer that the shorter the duration of the /aʊntɪŋ/ portion of a sample was, the shorter the /nt/ cluster was realized in it, along with a higher likelihood of flapping. As can be seen in this figure, the data are much more concentrated near the trendline compared to the word duration comparison. Therefore, at least based on the duration of surrounding segments, it can be said that hypothesis (3) is supported by the findings of the current data, meaning that a faster rate of speech indicates a higher likelihood of flapping intervocalic /nt/ clusters.

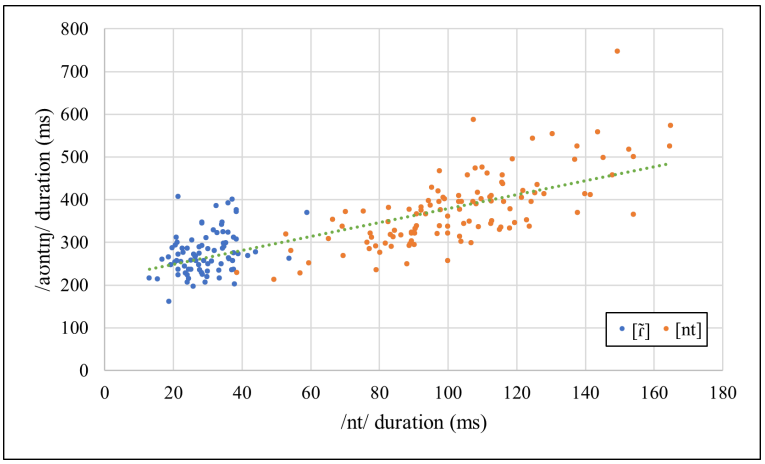

Figure 10: Scatterplot of the correlation between segment duration and the duration of the /nt/ portion of flapped and unflapped /nt/ samples

4 Conclusion

The current study focused on the flapping of /nt/ clusters, and its hypotheses were concerned with some of the factors which may influence the likelihood of flap production: lexical frequency, the gender of the speaker, and speech rate.

The first hypothesis was corroborated, as words with higher lexical frequency were generally more likely to exhibit flapped /nt/ clusters (although words with lower lexical frequency deviated from this trend), and the same was true when comparing utterances of the same word in various lexical categories. The second was supported by the data, which confirmed that male speakers were more likely to produce flaps than females, but the small sample size and unequal gender proportions do not provide a solid basis for this claim. The third hypothesis regarding speech rate was also verified based on the comparison of closure durations and word durations, meaning that the faster a speaker utters the word, the more likely they are to pronounce it with a flapped /nt/ cluster.

Future steps of the research include collecting samples from subjects in person, if a viable method can be designed, after which the tests presented in this paper can be rerun, other spectral features of the sampled data can be included in the research, and the role of word frequency and complexity, as well as gender, can be further inspected. It would also be beneficial to conduct some sort of perception experiment where listeners' ability to distinguish nasal flaps from /t, d/ flaps could be tested, and to perhaps explore other dialects of English for nasal flapping. The results of the present study can be useful to those wishing to construct new theoretical analyses of flapping, and they present empirical data and deductions relating to a previously (at least empirically) unexplored aspect of flapping which is evidently gaining ground in the everyday speech of Americans. Sound reduction processes are a natural element of every living language and they shape the way we use and interpret the language

in order to stay decipherable to fellow speakers and to be able to understand them as well. This is why it is important to stay up to date with the current state of linguistic phenomena such as flapping.

Acknowledgement

Supported by the ÚNKP-20-3 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 (2018). Audacity (version 2.3.3). https://www.audacityteam.org/.

Bermúdez-Otero, R. (2004). Raising and flapping in Canadian English: grammar and acquisition. Handout of paper presented at the CASTL Colloquium, University of Tromsø, 2 November 2004. www.bermudez-otero.com/tromsoe.pdf/.

Boersma, P. and Weenink, D. (2018). Praat: doing phonetics by computer (version 6.0.43). http://www.fon.hum.uva.nl/praat/. Accessed: 2021-04-05.

Braver, A. (2011). Incomplete neutralization in American English Flapping. *University of Pennsylvania Working Papers in Linguistics*, 17(1). https://repository.upenn.edu/cgi/viewcontent.cgi?referer=https://www.google.hu/httpsredir=1article=1165context=pwpl.

Davies, M. (2008). The Corpus of Contemporary American English (COCA): 600 million words, 1990-present. https://www.english-corpora.org/coca/. Accessed: 2020-06-15.

de Jong, K. J. (1998). Stress-related variation in the articulation of coda alveolar stops: Flapping revisited. *Journal of Phonetics*, 26(3):283–310.

de Jong, K. J. (2011). Flapping in American English. In van Oostendorp, M., Ewen, C. J., Hume, E., and Rice, K., editors, *The Blackwell Companion to Phonology*, pages 2711–2729. John Wiley Sons, Oxford.

Drinka, B. (1990). The chronology of phonological change. In Polom, E. C., editor, *Research guide on language change*, pages 107–136. Mouton de Gruyter, New York.

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

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

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

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

Kiparsky, P. (1982). *Explanation in Phonology*. Foris Publications, Dordrecht.

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech, Language, and Hearing Research*, 18(4):686–706.

Laver, J. (1994). *Principles of Phonetics*. Cambridge University Press, Cambridge, UK.

Patterson, D. and Connine, C. M. (2001). Variant frequency in flap production. *Phonetica*, 58(4):254–75.

Picard, M. (1997). English Flapping and the feature [vibrant]. *English Language and Linguistics*, 1(2):285–94.

Python Software Foundation (2019). Python (version 3.8.10). http://www.python.org.

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

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

The pandas development team (2021). pandas (version 1.2.4). https://pandas.pydata.org/.

Vaux, B. (2000). Flapping in English. Linguistic Society of America, Chicago, IL, 7 January 2000.

Zue, V. W. and Laferrière, M. (1979). Acoustic study of medial /t, d/ in American English. *The Journal of the Acoustical Society of America*, 66(4):1039–50.