Flapping in American English: A Theoretical Approach[1]

Julianna Sarolta Pándi

1. Introduction

In this paper I will discuss a phenomenon that is one of the most typical characteristics of American English, intervocalic alveolar flapping. Its occurrence can be observed in many languages besides English, both as an allophone of /t/ and /d/ and of /r/. It is also present in multiple dialects of English, such as Australian, New Zealand, Cockney, Irish, Scottish, South African and North American.

In my paper I am going to focus on its presence in American English, simply because that is the most widely covered occurrence in phonological literature. I will attempt to present a wide set of attributes, including but not limited to the possible analyses of flapping, its relationship to other allophones of /t/ and its effect on its environment.

Firstly, I will describe the basic rule of flapping, and examine the surface sound, the tap in Section 2. Section 3 is going to provide an elaboration of and explore the connection between tapping and flapping. Following that, I will discuss three parts of the rule: the input, the output and the environment in Sections 4, 5 and 6, respectively. In Sections 7 and 8 I am going to examine the possible effect of flapping on the preceding vowel. Finally, I am going to present an example and a possible theory in Section 9 for cases when flapping fails to occur in spite of the fact that rules that will have been discussed before are applicable to them.

2. The flapping rule

As stated above, intervocalic alveolar flapping is a phonological phenomenon occurring in many dialects of English. In my paper I will take a closer look at the occurrence of flapping in General American. Flapping is a rule stating that an intervocalic /t/ or /d/ surfaces as an alveolar flap [ɾ] before an unstressed vowel (Riehl, 2003).

The rule can be formalized as the following:

---

[1] I would like to take this opportunity to thank my supervisor, Péter Szigetvári, and my opponent, Miklós Törkenczy. All mistakes are exclusively mine.

$$\begin{Bmatrix} d \\ t \end{Bmatrix} \rightarrow ɾ \; / \; V\_V_0 \qquad V_0\text{: unstressed}$$

Figure (1): The flapping rule

Note that the rule above does not include all the possible environments of flapping, as it can occur after /n/ and /r/ and before syllabic /l/ as well, however, this paper only focuses on the intervocalic alveolar flapping phenomenon.

3. Tap vs. flap

The surface sound in the rule above may appear in two forms, a tap or a flap. The main line drawn between a flap and a tap (as described by Ladefoged & Maddieson, 2006:231) is that "a flap is a sound in which a brief contact between the articulators is made by moving the active articulator tangentially to the site of the contact, so that it strikes the upper surface of the vocal tract in passing; [whereas] a tap is a sound in which a brief contact between the articulators is made by moving the active articulator directly toward the roof of the mouth." In other words, the difference between the two features is the angle in which the tongue moves towards the passive articulatory organs.

Although they can be differentiated phonetically, there are no languages that contrast taps and flaps. Pointing out the one (or at least few) distinctive feature(s) between a tap and the /t/–/d/ alternants in which it is derived in flapping accents is hard enough on its own and there is no shared, consensual view on what said differences are. Making yet another distinction between these two surface sounds as well would complicate things unnecessarily. The fact that they are never in contrast in any known language explains why some phonologists are reluctant to distinguish between taps and flaps.

For some reason, the academic convention is to use the term "flapping" for the process but the term "tap" for the sound produced by it. In this paper I will follow these conventions and apply the same nomenclature.

In the next section, I will take a closer look at the flapping rule, examining the change itself, the input and the output, and the environment separately.

4. Archiphoneme

Firstly, let us consider the phonemes subject to flapping. The rule above states that the phonemes in question are the alveolar plosives, both the voiced /d/ and voiceless /t/. The only feature distinguishing them is voicedness, however when they undergo flapping this contrast

disappears, and consequently, it might become unclear which phoneme the tap represents. In other words, flapping neutralizes (suspends) a contrast otherwise present in the English language.

For instance, all native speakers are definitely able to tell whether the string of sounds they heard were of the forms *Ted* ([ted]) or *dead* ([ded]). In case of flapped forms, listeners have to rely a lot more on the context, for example with the string [ðə 'pɒɾ ɪz], as the phonological contrast between /t/ and /d/ (and consequently, the difference between the words *Ted* and *dead*) is suspended in this particular environment with flapping speakers.

It is interesting to note that due to Pre-Voiceless Vowel-Shortening; the vowels preceding /t/ are clipped compared to the ones preceding /d/. When the plosives undergo flapping, the consonantal distinction disappears, but vowel length difference prevails – this process is referred to as displacement of contrast. The possible perception of this contrast will be discussed later in Section 7.

Giegerich (1992) draws attention to the fact that the underlying forms of the phonemes whose contrast is suspended are not entirely arbitrary. As phonological rules always apply to a natural class (a set of sounds that have one or more phonetic features in common), it is safe to assume that the phoneme in question must belong to that particular natural class the rule targets. This way it is a logical option to categorise the input of these rules by their common features instead of simply listing them one by one (including all the features they differ in as well). This leads to a concept on a more abstract level than that of phonemes: that of an archiphoneme. The rule used for describing flapping could then be written as the following:

$$\begin{bmatrix} -continuant \\ -sonorant \\ +coronal \end{bmatrix} \rightarrow \text{ɾ} \; / \; [+sonorant] \_ \begin{bmatrix} +sonorant \\ -stressed \end{bmatrix}$$

Figure (2): The flapping rule using an archiphoneme

The input of this rule is an archiphoneme of an alveolar stop unspecified for voicedness, it can either be /t/ or /d/; therefore, the rule has the same targets of application as the one in (1).

5. Phonological and phonetic status of the tap

5.1. Phonetic properties

Even though its occurrence in General American is frequent and receives a lot of attention in phonological studies of the dialect, there are few things certain about the alveolar tap. Although its articulation is known fairly well (it is produced with a single contraction of the tongue muscles, which only makes a brief contact with the alveolar ridge), the question of what differentiates a tap from a stop is a controversial topic in the study of sounds. There have been some analyses supporting the idea of introducing a new abstract phonological feature to distinguish the tap from its stop counterparts. However, interpreting a feature for a few sounds can only lead to needless overgeneration in the feature system and, therefore, at least some evidence would be required to support that the newly established feature is present and needed elsewhere in the language as well, for which no evidence has been provided so far (Steriade, 1999).

Another candidate for possible distinction is the [+ sonorant] feature (Steriade, 1999). However, according to Hayes (2009:95) [t] and [d] are characterized by the features [−sonorant] [−continuant] and [−approximant], whereas the tap is [+sonorant] [+continuant] and [+approximant], from which it is obvious that only changing the binary feature of sonority would not instantly result in a [ɾ]. However, in Hayes's (2009:95) framework changing all three values to [+] would eliminate the contrast between [ɾ] and [r]. In Hayes the features introduced in order to still maintain the contrast are the tap and the trill ([ɾ] is described as [+tap] [−trill] and [r] as [−tap] [+trill]) but the introduction of two features raises the issue of redundancy[2]. Besides the inefficient use of features this is also troubling because from this representation the conclusion could be drawn that the tap is significantly more similar to [r] than to [t] and [d]. That does not reflect the behaviour of the sounds in General American, which is not necessarily a problem, but is also not an ideal condition.

[2]These features are relevant in the case of two other sounds: alveolar lateral tap [ɺ] and bilabial trill [ʙ]. However, the four sounds do not use the four-way distinction the two features would mathematically be able to make, the [ɺ] - [r] and the [ʙ] – [ɾ] pairs are distinguished on the basis of lateralness and labialness versus coronalness, respectively. This leads to a state in which the features [flap] and [trill] are not used to their maximal distinguishing potential.

It is worth pointing out that the alveolar tap's alternation with the stops is not a universal phenomenon. There are languages where the sound it alternates with is [r], for example Spanish and Dutch. According to Steriade (1999) "the tap [ɾ] is never in clear contrast, in the same system, with a homorganic voiced stop and a homorganic rhotic of identical moraic count."

The perceivable quantitative difference between the alveolar tap and the alveolar plosives can also account for their distinctness, unlike in the case of the two alveolar plosives[3]. Steriade (1999) introduced the feature of [extra short closure], which is, however, just one aspect that has been measured in isolation, disregarding the possibility of any other articulatory differences that might be present. Appointing [extra short closure] as the distinguishing feature between alveolar stops and alveolar taps would also, in a language lacking velar and labial taps, stand unsupported. The distinction would have to be suspended in the case of extra-short labial and velar stops, since there is no evidence supporting that their contrast exists at other places of articulation. In lack of the evidence from other sound changes any feature-appointment seems arbitrary. This theory will also be discussed in Section 9.

To sum up, we are yet to find a feature representation for the alveolar tap that fulfils all the criteria – avoiding a possible overgeneration in the system and representing the sound's relation to all the sounds it can be an allophone of, alveolar stops and the rhotic consonant.

5.2. Phonological properties

In this section I will examine the output of the flapping process, namely its status as a phoneme or an allophone (the aspects of this discussion below are based on McMahon, 2002).

For a sound to be considered an allophone of a phoneme, as opposed to a phoneme of its own, two criteria must be fulfilled. One is predictability, which means that we must be able to determine all the environments in which the sound will appear. A perfect example for a case in which the quality of a sound is completely predictable is that of clear and dark L's. It is well

[3] Based on the measurements of Zue and Laferriere (1979) and Banner-Inouye (1995) the difference between these sounds is measurable in milliseconds: an average tap is 26 ms, an average [d] 75 ms and an average [t] 129 ms long. Although the range of the absolute difference seems infinitesimal, it can be argued that proportionally it is huge: a tap is 4.96 times shorter than a [t] and 2.88 times shorter than a [d] which is considerably larger than the proportional difference sufficient to distinguish a [t] and a [d] (which is just 1.72).

known that in Standard British English an /l/ is realised with a velar secondary place of articulation if it is followed by a consonant or a pause and in all other cases without it. Therefore, in every environment it is completely predictable which allophone one is going to get (a clear [l] in 'kɪ_ ɪt' and a dark [ɬ] in the environments kɪ_ ðem or in kɪ_##). An instance of unpredictable difference is vowel backness in English. From the environment (no matter how broad) it is never predictable if a particular vowel is going to be front or back (one can never foresee whether the k_t environment will embrace an [æ] or an [ɔ] sound – i.e. whether the word in question is going to be *cat* or *cot*). Clearly, of these two, the tap is similar to our first example. It is entirely predictable where (within the speech of a speaker of a flapping dialect) it might occur, namely, exclusively in the environment mentioned in Section 2.

The other criterion to meet is that the sound in question must be able to express a meaningful contrast when compared to other allophones of the phoneme (i.e. the language cannot have words which can be distinguished by only these two allophones). The difference between [n] and [ŋ] is distinctive, a proof for which is the existence of the minimal pair /sɪŋ/ and /sɪn/. On the other hand, aspiration is a phenomenon that produces a non-distinctive sound: [p^h] will never be distinctive in English, one can never find two words with different meanings where one is with [p^h] and the other one is with [p], for instance. This condition is also fulfilled when analysing the tap — there are no minimal pairs in the English language that can only be distinguished by the opposition of [ɾ] and [t] or [d].

Based on the criteria discussed above we can conclude that the tap cannot be considered a phoneme in American English, but an allophone of /t/ and /d/. Although it is an allophone of both sounds, I will mainly focus on the relationship with /t/ as its numerous allophones present a wider range of variants for any further study.

5.3. The tap and other /t/-allophones

Allophones in the vast majority of cases are in complementary distribution, i.e. the environments they appear in can never overlap. That is not true of the alveolar tap – it is in free variation with the coronal plosives [t] or [d]. Either the flapped or the unflapped sound can appear in the environment given in Section 2 without causing any change in meaning. The ratio of their use is solely based on the speaker's idiolect and although different values may be attached to them socially, native speakers consider both variations equally grammatical.

Balogné Bérces (2005) describes /t/ as having the following eight different pronunciations in American English (based on Kenstowicz 1994: 65–66, Wells 1982: 248–252):

Allophone	**Transcription**	**Example**
"plain"	t	s<u>t</u>op
aspirated	t^h	<u>t</u>ake
retroflexed	ʈ	<u>t</u>rip
tapped/flapped	ɾ	ci<u>t</u>y
nasal flap	ɾ̃	pan<u>t</u>ing
unreleased glottalized	$t^ʔ$	hi<u>t</u>
glottal stop	ʔ	wha<u>t</u>ever
zero		pan<u>t</u>s

Figure (3): Allophones of /t/

In the following section I will examine the alveolar tap's relation to the glottalized [ʔt], the glottal stop [ʔ] and the aspirated [t^h]. According to Giegerich (1992) "[the tap] is in complementary distribution with [t^h], which occurs in onsets of stressed syllables, and with the [ʔt] of syllable-final positions." (For further discussion concerning syllable structure, see Section 6.)

Their complementary distribution alone (however relevant) is not enough to prove that for instance [t^h] and [ɾ] are the allophones of the same phoneme. The sounds [h] and [ŋ] are also in complementary distribution (while the former only occurs before vowels, the latter occurs generally before the consonants /k/ and /g/ or word-finally[4]). However, most analyses are reluctant to describe them as allophones of the same phoneme. The reason for that is quite simple: phonetically they are not similar at all (while [h] is a voiceless glottal fricative, [ŋ] is a velar nasal – and therefore voiced by default). In this case neither the place nor the manner of articulation is shared by the two sounds, let alone their voicing. In the case of the aspirated, the glottalized (or preglottalized) and the flapped allophones of /t/, however, the phonetic resemblance is obvious and striking.

[4] Except for a few words, e. g. *hanger*, *longer*, *stronger*, where the morpheme-medial velar nasal is in prevocalic position.

During aspiration the glottis is left open even after the voiceless obstruent is released, making the following sound devoiced. Glottalization means the complete closure of the glottis before the voiceless obstruent in question, which results in a clear-cut acoustic separation of the obstruent from the preceding vowel's phonation. Although these phonetic differences are present, none of the aforementioned modifications are distinctive.

All variants belong to the natural class of alveolar sounds and their manner of articulation is similar as well. This and the fact that they are incapable of expressing contrast (as a result of their complementary distribution) combined give enough ground for the assumption that the sounds mentioned above are in fact the allophones of the same phoneme, namely /t/.

To conclude, the tap sound is not a phoneme of English, but an allophone of /t/, and is in complementary distribution with other /t/-allophones, the aspirated and the glottalized /t/. While [t^h] occurs in both British and American dialects, [ʔt] is only typical of the former, while [ɾ] appears in the latter.

6. The syllable (determining the environment of the rule)

Finally, in this section, I will discuss the second part of the rule in (1), which describes the environment in which flapping occurs. The rule tells us that two things are needed for it to take place. The /t/ needs to be preceded by a vowel (of unspecified stress properties) and it also needs to be followed by an unstressed vowel. The phenomenon can be approached, however, from another point of view, one referring to syllable boundaries as well.

It is also a more desirable method to present aspiration, glottalization and flapping within the same framework, highlighting their similarities and unambiguously revealing their complementary distribution.

Recently syllables have mostly been thought of as possessing a structure that closely resembles the X-bar theory in syntax (Chomsky, 1970) as illustrated in Figure 4.

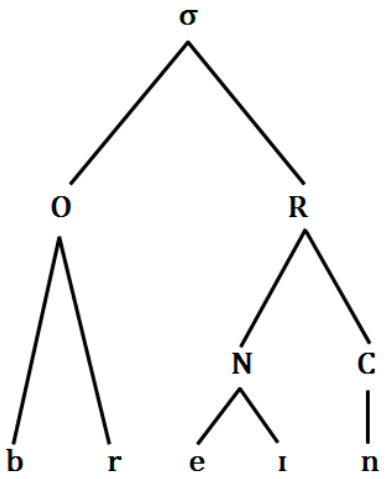

Figure (4): Syllable structure

In this hierarchic system, the syllable (σ) is divided into two parts first: the onset and the rhyme. The rhyme is further divided into a nucleus (or peak) and a coda[5]. The process of syllabification in English is governed by several rules. In the following I am going to elaborate on two of them: Onset Maximization and Coda Capture.

Onset Maximization is a tendency to have an onset consisting of as many consonants as possible. "Possible" means as many onset consonants as allowed by the sonority hierarchy. The sonority hierarchy is responsible for the fact that branching onsets are only possible if the second onset consonant is more sonorous than the one preceding it[6]. It is possible for a word that appears to be monosyllabic to contain two sonority peaks, for example a word-initial [s] within a consonant cluster, like in the word *stoic*. In this case the [s] forms a degenerate syllable; for further discussion please refer to Szigetvári (2011).

The sonority hierarchy is illustrated in Figure 5, based on Szigetvári (2012). Another factor (however much debated) that might limit the possibilities of a maximal onset is the set of possible word-initial clusters. It is proposed by Kahn (1980) that only clusters occurring at the beginning of words are able to occur at the beginning of a syllable. The logic behind this theory

[5] In most theories of moraic weight English rhyme components are perceived as being one mora heavy each (mora is the unit of weight in moraic theory). For further discussion on the Weight-by-Position rule see Hayes (1989).

[6]The same effect mirrored is true for the coda: each coda segment must be less sonorous than the one preceding it; therefore, the alternative name of the nucleus (peak) seems justified by it being the sonority peak of the syllable.

is that a word-initial cluster is necessarily an onset cluster, as it must form the onset of the syllable it is a part of.

Sonority index	**Sounds**
9	low vowels (e.g., a, ɑ, ɒ)
8	mid vowels (e.g., e, ɛ, ə, ɔ, o)
7	high vowels/glides (e.g., i, j, u, w)
6	rhotics (e.g., r, ɾ)
5	laterals (e.g., l)
4	nasals (e.g., m, n, ŋ)
3	voiced fricatives (e.g., v, ð, z, ʒ)
2	voiceless fricatives (e.g., f, θ, s, ʃ, x)
1	voiced plosives (e.g., b d g) (e.g., b, d, g)
0	voiceless plosives (e.g., p, t, k)

Figure (5): Sonority hierarchy

Coda capture (Kahn 1976) is a rule in syllabic theory that can possibly override the Onset Maximization Principle described above. It states that any strong syllable (i.e. a syllable with a major stress) is able to capture the onset of the following syllable as its own coda, provided that the following syllable is unstressed.

The result of Coda Capture is an ambisyllabic consonant (i.e. a consonant that cannot be unambiguously syllabified either as the coda of the preceding syllable or the onset of the following one). As syllabification does not necessarily coincide with morphological composition, Coda Capture (and therefore flapping) is not necessarily limited to the lexical domain either[7].

This theory, however debated, is perfectly suitable to explain the three-way alternation between glottalization, aspiration and flapping, which otherwise would be at least troublesome, since the syllable structure only offers two consonant positions.

Kahn (1976) reaches the conclusion that ambisyllabicity is a necessary requirement for flapping with the following arguments. First, he mentions that a /t/ or /d/ is only flapped if it is syllabified into the same syllable as the following vowel. By the rules of syllabification this is only possible if it forms a part of its onset. However, this leaves the onset position for both the flapping and the aspiration. We can easily identify the glottalizing environments by saying that it can only happen with coda /t/'s.

[7] A discussion on the relationship of morpheme boundaries and flapping at length can be read in Balogné Bérces (2005) and Kahn (1976).

Based on the fact that flapped /t/ is in complementary distribution with aspirated /t/, we know that their environments have to be different. Consequently, possible flapping environments are those which are both syllable-initial and syllable-final at the same time (i.e. ambisyllabic). From this it is easy to see that Coda Capture (CC) and flapping are necessarily related, namely, flapping applies to the output of CC.

To see it through an example, let's take the word *water*. Since the first syllable is stressed, coda capture is possible and it can be syllabified as *wat.er*, but this would mean that /t/'s not followed by a vowel can also be flapped. However, if we syllabify it as *wa.ter*, then the environment would coincide with aspiration and we still would not be able to clearly distinguish between the two phenomena. The theory of ambisyllabicity is able solve this problem, and also makes use of the rather rare situation in which both neighbours of the input sound are specified.

To conclude, the concept of ambisyllabicity provides a solution to restrictions of the two consonant positions within a syllable. Aspiration, as discussed above, happens to consonants in the onset of a syllable, for instance in words like *type* or *attack*. (Even though the /t/ is in an intervocalic position in *attack*, the first syllable is unstressed and thus unable to capture the coda consonant). Although the exemplary words were chosen in a fashion so as /t/ would be followed by a vowel, it is not necessarily so. Should /t/ be followed by non-nasal sonorants, aspiration would still occur, just not in the form of aspirated [t^h], but in the forms of a devoiced [r], as in the word [tr̥eɪ].

Glottalization or pre-glottalization happens in coda position, such as in the words *hat* and *what*. As previously reasoned, flapping will occur in ambisyllabic positions for which the examples are numerous in this paper. Thus, by using onsets, codas and ambisyllabicity, it is possible to represent three completely distinct environments, all in complementary distribution with the others.

7. Auditory cues (how we know it is there)

As it was already discussed in Section 4, there are measurable auditory cues that give indirect information about the voicedness of the original underlying plosive – namely the length of the vowels preceding the flapped stops. In spite of the fact that the difference between the plosives gets neutralized, the vowel difference still prevails – this process is referred to as a displacement of contrast.

However, it has been debated whether displacement of contrast eases perception. Numerous researches have come to the conclusion that vowels preceding the voiced plosive are on average longer than before its voiceless counterpart. The difference is not as drastic as between long and short vowels (it is only in the millisecond range) and results vary from study to study.

A study on this phenomenon was conducted by Herd et al. (2010), in which twenty American English speakers from Kansas participated. In the first part of the study all of them were asked to read a list containing words and nonwords contrasting /t/ and /d/ in flapping environments, both word-finally and word-medially. A total of 9840 tokens were recorded[8], and the acoustic nature of flapping (the duration of both the vowels and the taps) was analysed in detail.

For our discussion, the second part of the experiment carries the more relevant information. In this second part, a study of perception, 34 people participated. The researchers used 4 pairs of words, where a pair consisted of two identical surface representations, with one underlying /t/ and one underlying /d/ phoneme. The word pairs in question are: *leader–liter*, *wedding–wetting, tida–title* and *madder–matter*.

Their results have shown that native speakers are unable to identify acoustically for certain whether a sound in a given word is a flapped /t/ or a flapped /d/ — the pattern they have found is that the participants relied heavily on a d-bias (test subjects identified more /d/'s correctly than /t/'s and therefore, when in doubt, listeners were more likely to choose /d/).

Based on the aforementioned studies, we referred to the phenomenon of the vowel quantity before flapped sounds as a displacement of contrast. However, the paper by Herd et al. raises some interesting questions. From a perceptional point of view, the difference is fully suspended between the voiced and the voiceless plosives. If the basis of our analysis is only production, the difference prevails, but on a perceptual basis the contrast is suspended. This is an instance where the long-debated issue of phonological vs. phonetic representations recurs. The main problem here is how detailed a phonological representation is allowed to be and whether it is the articulation (the point of view of the speaker) or the perception (the point of view of the listener) that should be emphasized more and therefore represented in it. In this case the two clearly contradict each other, and although if one is thinking of practical usage, for

[8] For detailed distribution of the words please refer to Herd et al. (2010).

the creation of machines the two can be perfectly separable; for the purposes of writing a grammar which depicts exactly the relevant aspects of the natural language in question the decision is tougher.

8. Canadian Raising

In the next section, I am going to briefly digress from General American to examine a dialect where the neutralizing effect of flapping causes, without a doubt, a displacement of contrast: that is Canadian English (CE), and a phenomenon called Canadian Raising (CR). Moreover, I am going to discuss a possible analysis combining flapping and Canadian Raising.

According to Kaye (2012) Canadian Raising is the quality change of the diphthongs before voiceless consonants, /p/, /t/, /k/, /s/, /ʃ/ and /f/. Typically /aɪ/ becomes [ʌɪ], and /aʊ/ becomes [ʌʊ] – the output sounds may vary due to idiolects and regional dialects. What is common is that the [a]-component of the diphthong changes from low to mid-low, hence the name Canadian Raising. As a result, *price* is pronounced [prʌɪs] (as opposed to [praɪs] in General American), and *about* is pronounced [əbʌʊt] (as opposed to [əbaʊt]).

From our perspective, the words in Figure 6 present the relevant data, examining the word pairs *write–ride* and *writing–riding*.

	write	**ride**	**writer**	**rider**
GA	raɪt	raɪd	raɪɾər	raɪɾər
CA	rʌɪt	raɪd	rʌɪɾər	raɪɾər

Figure (6): Pronunciation of write and ride in GA and CE

In Canadian English, flapping and raising provide an example for displacement of contrast, as can be seen above in Figure 6. In a General American dialect, the words *writer* and *rider* are subject to the neutralizing effect of flapping – the only difference is present in the length of the preceding vowel. However, due to Canadian Raising, in CE the vowels differ in quality as well, and thus, we are able to phonologically distinguish them.

Through the example of these two phenomena, I am going to discuss rule-ordering, introduced by Kiparsky (1968). Whenever we have two rules that can possibly interact, we have to examine their relationship. CR can only occur before voiceless vowels, and that is exactly the feature flapping is able to delete. This is called a bleeding relationship – flapping is able to "bleed" Canadian Raising, in other words, is able to create a context where CR can no longer apply.

In the theory of rule-ordering, the idea is to determine how the rules follow each other in order to create the surface representations from the presumed underlying representations (UR). A model applying generative phonology would state that every word has an abstract underlying form stored in the mental lexicon of the speaker, and the words that actually appear in the language are all derived from the UR. By its very definition, we can see that the underlying form can never be more than an assumption, given its abstract nature. Consequently, every rule-ordering analysis includes a fundamental hypothesis – the underlying forms of each word. That is something that we cannot be sure of, and as I will discuss later, the question of UR is able to provide a foundation to challenge every rule-based theory. Keeping that in mind I am going to present this particular example by using the General American pronunciation as the underlying form.

In this case, since flapping is able to take away the environment necessary for Canadian Raising, we have to assume that CR occurs first and flapping second, creating a counterbleeding order. To test this hypothesis, in Figure 7 below I am going to examine the possible outputs for both cases. (7a) demonstrates the process if we assume flapping (FL) to occur first, and (7b) shows the order the other way round.

	write	**ride**	**writer**	**rider**
UR	raɪt	raɪd	raɪtər	raɪtər
FL	raɪt	raɪd	raɪɾər	raɪɾər
CR	rʌɪt	raɪd	raɪ**ɾ**ər	raɪ**ɾ**ər

Figure (7a): Rule-ordering in Canadian English I.

The forms in bold above show the reason why this order cannot be the right one – if we assume flapping precedes Canadian Raising, the neutralizing effect would still prevail, and the forms surfacing would not reflect the words of the language. That is a direct cause of flapping's bleeding nature in this order. However, if CR is the first to occur, the order prevents flapping to take away the necessary environment, and the derived surface representations will be the correct ones.

	write	**ride**	**writer**	**rider**
UR	raɪt	raɪd	raɪtər	raɪtər
CR	rʌɪt	raɪd	rʌɪtər	raɪdər
FL	rʌɪt	raɪd	rʌɪɾər	raɪɾər

Figure (7b): Rule-ordering in Canadian English II.

As the examples above clearly show, flapping and Canadian Raising are an obvious example for demonstrating counterbleeding and rule-ordering. However, Kaye (2012) in his research challenges this idea based on "the assumption (without supporting evidence) that *write* and *ride* are derived from the same underlying vowel" among others.

To sum up, in the section above I presented an example of Canadian English to show a dialect in which flapping causes an obvious displacement of contrast. By using these two rules, we are able to demonstrate rule-ordering – specifically, a counterbleeding relationship between the two rules.

9. Paradigm Uniformity

Based on the attributes discussed above, flapping appears to be a post-lexical rule. It applies across word boundaries (e.g. ['gɔ ɾ ɪt]), it is optional and not structure-preserving — it produces a sound which is not part of the phoneme inventory of the language.

However, there are exceptions, which is not typical of post-lexical rules — in this section I am going to discuss an environment where flapping fails to occur, and a possible explanation offered by Steriade (1999).

The phenomenon in question is called the Withgott-effect. Withgott (1983) noted the different pronunciation of the stop in *capitalistic* and words such as *militaristic, sanitisation, monotonocity* – in the former, the /t/ is flapped and in the latter ones it is not. The difference cannot be attributed to the CV patterns or the stress pattern, since each word shares the same CV pattern and stress distribution. However, the /t/ in the base forms of the words *capital* and *military* are not the same, as only the former is flapped.

Steriade (1999) attributes the phenomena to Paradigm Uniformity (PU). A paradigm is a set of words that share a morpheme (either root or suffix) and Paradigm Uniformity is a theory that states that the morphemes have a tendency to retain their original form throughout the paradigm. The process the morphemes go through is called Paradigm Levelling, when a speaker uses an unexpected or unjustified form to fulfil the requirements of PU. Steriade uses

the term *phonetic analogy* when describing Paradigm Levelling resulting in the change of segments that are not known to be independently contrastive in any language.

The following experiment by Steriade (1999) can be used to illustrate PU. Native speakers were asked to produce base words with antepenultimate stress (*jettison*, *parody*), and then to produce the derivative forms suffixed with *-able*, creating nonces. The reason why this particular suffix has been chosen is that by adding *-able* the stress pattern of words typically remains intact. However, leaving the stress where it was in the base form violates a prosodic constraint of English which generally does not permit unstressed strings longer than two syllables.

The vast majority of speakers preferred the forms in which the stress pattern of the base word was retained, for example [ˈdʒɛtɪsənəbəl] to [dʒɛtɪˈsɒnəbəl]. Although they described the resulting words ‘too long’ and ‘awkward’, they stated that the pronunciation where the stress has been moved is unacceptable. From this it is clear that native speaker intuition recognises that in these cases there are two opposing phonological tendencies that have to be considered, and a majority of speakers feel that retaining Paradigm Uniformity is the best option the system is capable of. It is important to note that PU is not a rule that must be obeyed under all circumstances; the linguistic judgement can even vary from speaker to speaker.

Moreover, the length preservation of vowels preceding taps (the phenomenon described in Section 7 at length) can also be considered as an example for PU. According to Steriade (1999), the reason for the unflapped stop in *militaristic* is the same tendency, namely, that speakers prefer to retain the consonant that was present in the base form, *military*. By moving the stress, Paradigm Uniformity has already been violated; therefore it is even more important to retain the similarity of the surface sounds. Even though flapping is not a distinctive feature and thus does not affect comprehensibility, the surface form of the morpheme would still change.

In order to support her claim, Steriade conducted an experiment in which 12 subjects were asked to read four lists of words (as presented in (8)): two sets of base forms and two sets of derived forms suffixed with *-istic*.

(8a) Bases: *voluntary, positive, primitive, relative, negative*
Derivatives*: voluntaristic, positivistic, primitivistic, relativistic, negativistic*

(8b) Bases: *rotary, fatal, fetish, totem, notary*
Derivatives: *rotaristic, fatalistic, fetishistic, totemistic, notaristic*

Figure (8): Word pairs used in production study (Steriade, 1999)

The words in (8a) are the ones the experiment actually focused on – their flapping is either optional or impossible. The ones in (8b) are all potential subjects to flapping, they were included to avoid false positive [t]'s. If the list included words that are almost certainly flapped (since the whole phenomenon is optional, we cannot predict flapping with absolute certainty), the subjects are less likely to feel pressured to overproduce [t]'s. Every subject was asked to read two lists of words, composed of the four different sets presented above, first a list composed of the base forms, then after a short break one of the derivative forms.

The results have shown that the base forms in (8a) and (8b) did not have any effect on one another, all base forms in (8b) were flapped, except for one speaker who did not tap *fetish*. Whereas all the /t/'s of the derived forms in (8b) surfaced as flaps, the ones in (8a) varied. The flapping/non-flapping ratios were the following: 1/5 for *positive*, 7/5 for *primitive* and *relative*, 2/1 for *negative*. In the case of base form-derivative form pairs 11 speakers' results showed complete correspondence throughout the paradigm.

The research was criticized by Riehl (2003) on multiple counts. She drew up a few possible problems regarding the experimental method. Firstly, she states that if every subject says every word only once, we cannot be sure that the results are caused by Paradigm Uniformity, and not by sheer coincidence. She also claims that the perception of taps and stops in the produced words have to be added to the study as a second part, since in the original study by Steriade the choice between taps and stops was only up to her decision. According to Riehl, the judgement of the researcher is not a sufficient indicator.

In the light of these results, Steriade (1999) studies what feature distinguishes taps from stops and arrives at the conclusion that it is the [extra short closure], mentioned in Section 5.1. This notion is also criticized by Riehl (2003), on the basis of the following: although Steriade (1999) states that the difference in closure length is measurable, "she does not subject the tokens in her study to an acoustic analysis and therefore is unable to verify whether the tokens

she perceives as flaps are actually shorter in duration than the tokens she perceives as stops, or whether other cues relate to flap/stop identification as well as or better than duration."

Based on her experiment, Steriade proposes the idea of obliterating the distinction between phonetic and phonological features. She challenges Keating's (1984) concept of phonological features. She claims that only those features are validly considered phonological that are able to express contrast by themselves. However, Steriade (1999) offers a new perspective by stating that Keating's features are in fact phonological contrasts, which are perfectly analysable as the totality of several phonetic features. For example the contrast in the voicing of consonants stems from features such as closure duration, prevoicing, pitch, etc. Since these features on their own are non-contrastive, they are regarded phonologically irrelevant and changing only one of them is incapable of sufficiently expressing contrast.

According to Steriade (1999), the [extra short closure] belongs to the category of non-contrastive features. However, it is able to trigger Paradigm Levelling, and consequently, is sufficient proof for non-contrastive features to be considered as grammatically relevant.

If we accept Steriade's stipulation that the difference between stops and flaps is only the closure length, her conclusion is absolutely logical.

This difference cannot be considered strictly phonetic, because if it were, listeners would not be able to differentiate between the two sounds, since Steriade herself claims that one phonetic feature is not sufficient on its own to express contrast. However, it cannot be considered as phonological either. There is only this one example for the phenomenon (as mentioned in Section 5.1, no such distinction is made among labials or velars), and even in this case its presence is predictable. This phenomenon by itself does not provide sufficient evidence against predictability as an indicator of phonological irrelevance.

She resolves this opposition by rendering the distinction between phonetic and phonological features unnecessary. However, in my opinion, this theory only stands as long as we accept the [extra short closure] as the only distinction, and as Riehl (2003) stated, we do not have satisfactory evidence to support that.

While I do not disagree with the final conclusion Steriade (1999) arrives to, namely that distinction between phonetic and phonological features should be reconsidered, I do feel that the alveolar flapping is not the ideal phenomenon to support the idea.

10. Summary

In this paper I aimed at presenting a detailed description of the intervocalic alveolar flapping in American English. I examined the input, the output and the environment of the rule. The input is a good example to support introducing the concept of the archiphoneme, an abstract phoneme unspecified for redundant features. The output of the rule, the tap itself, can be discussed from a phonetic and a phonological point of view – the features that distinguish it from /t/ and its allophony to /t/, respectively. And finally, the environment presented a notion in syllabic theory, ambisyllabicity, which states that a consonant is able to belong to multiple parts of a syllable.

In the following sections, I discussed the environment of flapping. Multiple studies have been carried out to examine the length of the preceding vowel, however, in my research I have not found a framework that is able to combine the perceptional and articulatory results into a single, unified rule set.

After that, I presented a dialect which shows a change in quality of the preceding vowels, Canadian English, and I gave an example of rule-ordering by using flapping and Canadian Raising. And lastly, I examined the Withgott-effect, the cases where flapping does not occur, a possible explanation for that by Steriade (1999) and its critique by Riehl (2003).

Throughout this paper a few problems were presented that could be subject to future research. Firstly, the phonetic representation of the tap is far from being trivial; there is no universally accepted analysis to describe the features of the sound. Secondly, while the notion of ambisyllabicity perfectly solves the problem of the three-way alternation, it is not a widely accepted theory, and it has been debated whether it really is needed in English phonology. Third is the question of production versus perception. With the example of flapping (more precisely, the vowels preceding the flapped sounds) we can address an interesting issue, a question of what the best basis of a phonological model would be: the sounds articulated or the sounds perceived by the speakers. And lastly, the issue raised by Steriade (1999) was also presented through the example of flapping, namely the need to redefine the distinction of phonetic and phonological features. While her study was challenged by other phonologists, and I claimed that the example she chooses does not seem to support her argument, I do think that the initial idea is worthy of thorough research.

References

Balogné Bérces, Katalin. 2005. *Strict CV phonology and the English cross-word puzzle.* Saarbrücken: VDM Verlag Dr. Müller Aktiengesellschaft & Co.

Banner-Inouye, Susan. 1995. *Trills, Taps and Stops in Contrast and Variation.* PhD dissertation, UCLA.

Chomsky, Noam. 1970. Remarks on nominalization. In: Roderick Jacobs and Peter Rosenbaum (eds.) *Reading in English Transformational Grammar.* Waltham: Ginn. 184-221.

Giegerich, Heinz J. 1992. *English Phonology.* Cambridge, UK: Cambridge University Press.

Hayes, Bruce. 1989. *Compensatory Lengthening in Moraic Phonology.* Linguistic Inquiry, 20: 253-306.

Hayes, Bruce. 2009. *Introductory Phonology.* Chichester: Wiley-Blackwell.

Herd, Wendy, Allard Jongman, and Joan Sereno. 2010. An acoustic and perceptual analysis of /t/ and /d/ flaps in American English. *Journal of Phonetics*, 38: 504–516.

Kahn, Daniel. 1980. *Syllable-based generalizations in English phonology.* New York: Garland.

Kaye, Jonathan. 2012. Canadian Raising, eh? In: Eugeniusz Cyran, Henryk Kardela, and Bogdan Szymanek (eds.) *Sound, Structure and Sense: Studies in Memory of Edmund Gussmann.* Lublin: Wydawnictwo KUL. 321-352.

Keating, Patricia A. 1984. Phonetic and phonological representation of stop consonant voicing. *Language*, 60: 286-319.

Kenstowicz, Michael. 1994. *Phonology in Generative grammar.* Cambridge, MA: Blackwell.

Kiparsky, Paul. 1968. Linguistic Universals and Linguistic Change. In: Emmon Bach and Robert Harms (eds.) *Universals in Linguistic Theory.* New York: Holt Rinehart and Winston. 171-202.

Ladefoged, Peter and Ian Maddieson. 2006. *The Sounds of the World's Languages.* Oxford: Blackwell.

McMahon, April. 2002. *An Introduction to English Phonology.* Edinburgh: Edinburgh University Press.

Riehl, Anastasia K. 2003. American English Flapping: Perceptual and Acoustic Evidence Against Paradigm Uniformity with Phonetic Features. *Working Papers of the Cornell Phonetics Laboratory*, 15: 271-337.

Steriade, Donca. 2000. Paradigm Uniformity and the phonetics-phonology boundary. In: Michael Broe and Janet Pierrehumbert (eds.) *Papers in Laboratory Phonology V: Acquisition and the Lexicon*. Cambridge, UK: Cambridge University Press. 313-334.

Szigetvári, Péter. 2011. The syllable. In: Bert Botma, Nancy C. Kula, and Kuniya Nasukawa (eds.) *Continuum Companion to Phonology*. London & New York: Continuum. 64-94.

Szigetvári, Péter. 2012. Phonological analysis 4: Syllable structure. In: Zoltán G. Kiss, Ádám Nádasdy, Attila Starčević, Péter Szigetvári, and Miklós Törkenczy. 2012. *Phonological analysis.* Budapest: Eötvös Loránd University. (http://seas3.elte.hu/analysis/book.html)

Wells, John C. 1982. *Accents of English.* Cambridge, UK: The Press Syndicate of the University of Cambridge.

Withgott, Mary. 1983. *Segmental evidence for phonological constituents.* PhD dissertation, University of Texas, Austin.

Zue, Victor W. and Martha Laferriere. 1979. Acoustic study of medial /t, d/ in American English. *Journal of the Acoustical Society of America*, 66.4: 1039-1050.