On The Phonologization of Partially Contrastive Phones /t/ and /tʃ/ in American English via Two Experiments

Undergraduate Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation *with honors research distinction* in Linguistics in the undergraduate colleges of The Ohio State University

By

Margot Hare

The Ohio State University

April 2024

Project Advisor: Dr. Rebecca Morley, Department of Linguistics

Abstract

This thesis explores the phenomenon of /t/ affrication before /ɹ/ in American English. Over the past several decades, linguists have observed that [t] has become more often pronounced as [tʃ] when preceding /ɹ/ in general English speech, in such words as ‘tree’ [tʃɹi] and ‘truck’ [tʃɹʌk]. This change was likely spurred by coarticulation, a phonetic phenomenon where a sound unit is affected by the physical characteristics of a nearby sound. In this case, the further-back-in-the-mouth /ɹ/ triggers the /t/ to be produced further back in the mouth. This is an example of retraction, a type of coarticulation. Because a retracted /t/ becomes more like the affricate /tʃ/, it may also be termed “affrication”. We tested the hypothesis that this phenomenon has been phonologization using two experiments. In the perception experiment, we investigated how /t/ before /ɹ/ is perceived and categorized as a function of the phonetic factors speaking rate, word location in a sentence, syllable count, and vowel. Next, we conducted a corpus study to analyze the change over time of the phenomenon of /t/ affrication before /ɹ/ in Columbus speakers and observed speaking rate’s effects on this affrication. The effects of the phonetic factors were used as a metric to discern whether the affrication phenomenon has undergone phonologization.

The results of the perception experiment fail to provide clear evidence for or against the phonologization conclusion, with results pointing in both directions, although, evidence may be leaning against phonologization. The results of the corpus study indicate that the phonologization conclusion should be rejected. However, no consensus was found, and, as such, the phonologization question is still open to investigation.

Introduction

This paper is primarily concerned with the distinction between phonological processes and phonetic processes and into which of these two categories the /t/ affrication before /ɹ/ phenomenon falls. Phonology as a linguistic field focuses on the abstract, mental representations of sounds and how sound systems are organized in human language. Traditionally, slashes are used to represent sounds as they are stored in one's head (e.g., /t/ if the speaker intends to produce a /t/). Phonetics refers to the physical aspects of sounds, such as the physical articulations required to produce a sound and a sound's physical characteristics. When referencing the actual physical production of a sound, brackets are used (e.g., [t] when a true [t] is produced). Phonological processes are traditionally viewed as categorical, where sound features like [+nasal] either spread or do not, as opposed to phonetic processes where the degree of sound characteristic spreading may be gradient (different degree, e.g., duration, of nasalization) as a function of phonetic environment (Bermúdez-Otero & Trousdale, 2012; Kilbourn-Ceron & Sonderegger, 2018). Over time, phonetically motivated sound changes may become phonological, rather than phonetic, in nature (Bermúdez-Otero & Trousdale, 2012; Hyman, 2013). This is termed "phonologization". One phonetic process that can, but does not necessarily, lead to this sound change is coarticulation, whereby a sound changes gradiently to become more phonetically similar to nearby phones. It is unclear if the /t/ affrication before /ɹ/ phenomenon is still a result of coarticulation or if it has moved beyond into phonology.

In the following sections, I discuss how sounds can be related to each other, exploring the different relationship types in order to illustrate the uncertainty in how the sounds /t/ and /tʃ/ relate to each other (*Relationships Between Phones*). Next, I discuss how the type of relationship shared between a pair of phones influences how those phones are perceived in relation to each other

(*Relationship Effects on Perception*) and how the differing relationships in different phonetic environments can result in the phonetic environment affecting how phones are perceived (*Phonetic Environment Effects on Perception*). This leads to a pilot study I conducted on this topic which spurred further investigation. The term "phonologization" is then further fleshed out and explained (*Phonologization*), and, finally, I discuss the direct experimental motivation and inspiration for the current study and the reasoning behind its methodology (*Methodological Foundation, Experimental Reasoning*).

Relationships Between Phones

In phonology, pairs of phones, i.e. individual speech sounds, can have various relationships with each other. These relationships include contrastive and non-contrastive. Contrastive pairs are two phones that can occur in the same phonetic environment and change the meaning of a word when substituted for each other, creating a contrast. Such word pairs are called minimal pairs, where the words are minimally different from each other because they differ by only one phone. For example, in English, the phones /f/ and /v/ constitute a contrastive pair because of minimal pairs like 'fat' [fæt] and 'vat' [væt]. Swapping one sound for the other completely changes the word's meaning.

Non-contrastiveness is another relationship type. Hume and Johnson (2003) suggest that the non-contrastive relationship actually encompasses multiple types of non-contrastiveness. One type is allophony, or a complementary distribution. Allophonic pairs are pairs of phones that do not create a contrast in meaning because they are "distributed" (appear) based on their surrounding phonetic environment and do not appear in the same environment. For example, in English, the aspirated /pʰ/ occurs exclusively at the beginning of stressed syllables (the conditioning phonetic environment), such as in 'pin' [pʰɪn], while the unaspirated /p/ appears elsewhere, as in 'spin'

[spɪn]. Thus, /pʰ/ and /p/ are in a complementary distribution and allophonic relationship. Allophonic relationships can be either phonological or phonetic. Another type of non-contrastiveness is partial contrast (Hume & Johnson, 2003), or "context-dependent contrastiveness", as termed by Celata (2007). This occurs when a pair of generally contrastive phones are "neutralized" in a specific context, i.e., one of the phones appears and the other does not (no contrast). Thus, two phones may have different relationships with each other depending on the phonetic context in which they appear. Hume and Johnson (2003) present the example of the low-falling-rising and mid-rising tones in Mandarin Chinese which are contrastive except when following a low-falling-rising tone. In this case, only the mid-rising tone appears.

However, there are yet still more types of non-contrastiveness that do not fall cleanly into either of these categories. This is true of the phone pair under investigation in this study: /t/ and /tʃ/ (typically written as "t" and "ch", respectively, in American English). They are neither fully contrastive nor fully non-contrastive, and they are not clearly allophonic or in partial contrast. They are contrastive in every word position: word initial, as in 'tart' [tɑɹt] and 'chart' [tʃɑɹt]; word final, as in 'teat' [tit] and 'teach' [titʃ]; and word medial, as in 'lector' [lektɝ] and 'lecture' [lektʃɝ]. Swapping the phones in each of these word pairs produces a change in meaning (contrastive relationship). On the other hand, in words like 'tree' and 'attractive', one can produce the "t" sound as /t/ or /tʃ/ ('tree' [tɹi] vs. [tʃɹi], 'attractive' [ətɹæktɪv] vs. [ətʃɹæktɪv]), and no change in meaning occurs (the appearance of the plosive sound /t/ as the affricate sound /tʃ/ is termed "affrication"). Because of the inconsistent appearance of /tʃ/, this could be an example of free variation, where two (or more) phones can be interchanged in the same word by the same talker without changing the meaning. However, the concept of free variation is somewhat controversial (is it really free? does it even exist?) and is not further explored here (Anttila, 2018). It may also be argued that this

is an example of an allophonic relationship, motivated by the phonetic environment, where the following /ɹ/ conditions the appearance of [tʃ] instead of [t]. However, this has not been directly experimentally supported, and the existence of an allophonic relationship between /t/ and /tʃ/, whether phonetic or phonological, is not under investigation here. Alternatively, this phenomenon of [tʃ] appearance instead of /t/ before /ɹ/ may be in an in-between state, still undergoing change in the phonological system of American English. We explore this idea in this work.

Relationship Effects on Perception

Because the current study employs a perception study, the discussion of how relationships affect perception is of particular relevance. In 2008, Boomershine et al. produced a paper investigating how the phonological relationships between phones in one's native language affect one's perception of those sounds. This paper was motivated by previous literature, which suggested that if two phones are not contrastive in one's native language, they are harder to differentiate (Goto, 1971; MacKain et al., 1981). Boomershine et al. explored the phones /d/ and /ð/. These two sounds are contrastive in English but allophonic in Spanish. For example, English has the minimal pair 'dare' [dɛɹ] and 'there' [ðɛɹ]. Boomershine et al. presented 'de [ð]onde' 'from where' and '[d]onde' 'where' as an example of [d]/[ð] allophony in Spanish, where [d] occurs at the beginning of phrases and after /n/ and /l/ and [ð] occurs elsewhere. This study found that English speakers rated the two sounds as being more different, and Spanish speakers rated them as more similar. Thus, relationships between sounds in our native language, i.e., our phonological system, affect how we perceive these sounds. Furthermore, in 2003, Hume and Johnson found that "perceptual distinctiveness is a function of phonological contrast and that partial contrast reduces perceptual distinctiveness for native listeners", meaning that two phones that are neither fully contrastive nor fully non-contrastive are perceived as being less distinct.

Phonetic Environment Effects on Perception

Scudieri (2012) took this idea of relationship affecting perception a step further with a specific partially contrastive phone pair, /s/ and /ʃ/, by incorporating phonetic environment into the equation. Like /t/ and /tʃ/, the phones /s/ and /ʃ/ can be swapped in word pairs like 'fasten' [fæsᵊn] and 'fashion' [fæʃᵊn], and it creates a change in meaning. However, in a word like 'street', where /s/ is followed by /tɹ/, the first sound can be produced as /s/ or /ʃ/ ([stɹit] vs. [ʃtɹit]), and the meaning stays the same. Thus, the sounds are non-contrastive in that environment and may be perceived as less distinct. Scudieri provided evidence that the physical phonetic context in which a phone is located can affect a person's perception of the phone. She did this by manipulating the phonetic environments of /s/ and /ʃ/ by either inserting or not inserting [tɹ] after it and found that the word pair /ʃtɹɑ/ and /stɹɑ/ (with inserted /tɹ/) was rated by participants as more similar than the pair /ʃɑ/ and /sɑ/ (without /tɹ/) where still the only difference between the words in each pair was [ʃ] vs. [s].

Thus, the unclear relationship between /t/ and /tʃ/ in words like 'tree' may be elucidated by exploring their perception in different environments. In preparation for the current study, we conducted a pilot study in Spring 2023 to investigate /t/ and /tʃ/ perception in different environments. In this study, participants listened to pairs of made-up words that were recorded by the researchers. They were instructed to decide whether the two words were the same or not (an AX Discrimination Task). Similarly to the Scudieri study, the word pairs of interest differed only by 1) the initial consonant: [t] or [tʃ] and 2) the presence of /ɹ/ following the initial consonant. The goal of this pilot was to determine if the phonetic environment of having an /ɹ/ or not affected the perceived similarity of the two consonants. As predicted, the participants incorrectly judged [tɹ] vs [tʃɹ] initial word pairs as being the same 80% of the time (the consonants were perceived as more similar). However, there was enough variability in responses among the participants, with a

few participants judging the words to be different most of the time, to indicate that the relationship between the sounds may not (yet) be the same for everyone. There may be other factors at play other than simply the following /ɹ/ environment.

Phonologization

The goal of the current study was to determine if this sound pattern, /t/ → [tʃ] before /ɹ/, has become "phonologized". While this term has historically been nebulous (Kiparsky, 2013; Yu, 2021), most discussions revolving around phonologization eventually fall back on Hyman's definition, referring to phonologization as "the change of a phonetic property into a phonological one" (2013, p. 4). Another way to think of this is that some phonetic constraint undergoing sound change (e.g., coarticulation) is exaggerated until it becomes noticeable and goes beyond the phonetic level (Hyman, 2013). At this point, it becomes a phonological, not phonetic, process whereby a phone changes categorically to match a category of a nearby phone. In this way, a new articulatory target is adopted.

Methodological Foundation

Magloughlin (2018) similarly examined the question of the phonologization of the [t] vs. [tʃ] phenomenon, and the current study is somewhat a replication of their work with alterations. Magloughlin remarked that, while this phenomenon is often taken for granted in the literature as being true, there have been few if any experimental investigations into it. Magloughlin bases the identification of phonologization on the change to a new phonological articulatory target motivated by initial coarticulation. In this case, the coarticulation is the further-back-in-the-mouth /ɹ/ triggering the /t/ to be produced further back in the mouth. First, Magloughlin conducted a corpus study, which is a type of study that involves the analysis of a large collection of natural text or spoken language samples (a corpus). In this study, Magloughlin examined the speech of Raleigh,

North Carolina natives from varying age groups (birth years ranged from 1923-1996, split into three groups). This work was undertaken to determine if this phenomenon was indeed a sound change in progress or the effect of sudden contact with other dialectal groups moving into the Raleigh area. Magloughlin found that, as speaker age decreased, the /tɹ/ pronunciation became less like [t] and more like [tʃ], suggesting a gradual sound change of increasing affrication over time beginning in the mid-twentieth century in Raleigh.

Magloughlin also conducted a perception study, utilizing a forced-choice spelling task, with the goal of determining how the affricated /t/ is categorized by listeners (as /t/ or /tʃ/). Participants were presented with the recording of a made-up word and given two options of how to spell it. The spelling choices differed only in the initial letter(s); one started with "t" and the other with "ch". Each recorded word consisted of two parts: an initial consonant sound and either a vowel-initial sequence or an /ɹ/-initial sequence. The initial consonant sounds were either 1) /tʃ/ spliced out of a recording where it appeared before a vowel, 2) /t/ spliced out of a recording where it appeared before a vowel, or 3) /t/ spliced out of a recording where it appeared before /ɹ/. The idea was that the /t/ spliced from before /ɹ/ would be more affricated. The results showed that participants overwhelmingly spelled /t/ from before /ɹ/ using "t" when it was followed by an /ɹ/, but they spelled /t/ from before /ɹ/ with "ch" when followed by a vowel. This reveals that a pre-/ɹ/ /t/ is affricated enough to the point that is it no longer recognized as /t/, but rather as /tʃ/, when not in that conditioning environment. This exhibits perceptual compensation on the part of the listeners. These findings, coupled with observations of conscious awareness of this effect (e.g., children spelling 'tree' as "chree"), support Magloughlin's hypothesis that the [tʃ] variant before /ɹ/ coarticulatory effect has been phonologized. The results of our pilot study, however, urged us to take a second look.

Experimental Reasoning

Magloughlin's work forms the foundation for the methodology of the current study, which asks the same phonologization question with an altered methodological approach. The approach used in Magloughlin's study – changing the environment via splicing – may not agree with our approach: changing phonetic factors. Two experiments were conducted in this study: a perception experiment exploring the phonetic effects of speaking rate, word location, syllable count, and vowel on perception (measured via spelling choice of "t" or "ch") and a corpus study exploring the impact of gender, age, and speaking rate on affrication of /t/ before /ɹ/.

Speaking rate, which influences speech on a phonetic level, has been shown to impact listeners' perceptions, including the perception of certain phone and stress distinctions and even the appearance of certain words (Dilley & Pitt, 2010; Reinisch et al., 2011; Summerfield, 1981). It has also been shown to affect the rate of coarticulation, which is a phonetic process (Agwuele et al., 2009; Gay et al., 1974). Word location, syllable count, and vowel are other phonetic variables that may produce similar effects.

Because phonological processes are traditionally viewed as being categorical, where they do or do not happen, we would not expect to see an effect of these phonetic influences if the phenomenon has been phonologized (Phillips, 2018). Thus, in the opposite direction, if the affrication phenomenon has not been phonologized, we would expect more coarticulation (resulting in affrication) at a faster speaking rate and less at slower speaking rates. Finding an effect of speaking rate on production in the corpus study would be evidence against phonologization.

In terms of perception, if this is a coarticulatory (phonetic) effect and not phonological, listeners may "expect" to hear more affrication at faster speaking rates and perceive an affricated

/t/ in a fast sentence as being underlyingly a /t/ that was coarticulated. Affricated /t/'s at slower speaking rates would be less expected to have undergone this coarticulatory process, resulting in the affricated /t/ being interpreted as a true /tʃ/ underlyingly. In this case, we would expect to see higher rates of participants spelling /tr/ words with "t" at higher speaking rates and lower rates of participants spelling /tr/ words with "t" (conversely, higher "ch" spellings) at lower speaking rates. Effects from word location, syllable count, and vowel are harder to predict as their connection to articulation is less direct. However, similarly, finding an effect of these phonetic variables on perception would be evidence against phonologization.

Two Experiments

Perception Experiment

This experiment used a forced choice spelling task methodology to investigate the effect of the independent variables of speaking rate, word location in a sentence, syllable count, and vowel on spelling choice.

Methodology

Participants sat at a computer in a soundproof booth. The participants listened, using headphones, to recordings of made-up words embedded in carrier sentences of differing speaking rates. Embedding the words in carrier sentences allowed the listeners to infer the speaking rate. Speaking rate was divided into two categories: slow and fast. There were 32 total sentences. 16 sentences were spoken at a "fast" rate, and 16 were spoken at a "slow" rate. While listening to a sentence, participants were shown the sentence written on the screen with a blank where the non-word was. For example, if a participant heard the sentence, "Let's discuss mrih," they would see "Let's discuss ____," written on the screen. After hearing and reading the sentences, the participants were presented with two options on the screen of how to spell the made-up word they

had just heard in the given sentence. For example, “mrih” or “brih”. To make their spelling decision, the participants pressed one of two buttons on the keyboard.

Stimuli

The primary researcher recorded 32 sentences to be used as stimuli in the soundproof booth. 16 of these 32 sentences contained filler words that were not under test and were included only to provide more variation for the participant. These were words like “brih”, “ahmrah”, and “uhbruh”. 16 of the sentences contained the made-up words of interest, i.e., words containing [tɪ]. All words were inserted into sentences where it was clear that they were nouns in order to avoid any affect from grammatical variation. In each of these two groups (filler words vs words of interest), 8 sentences were spoken at a “fast” rate, and 8 were spoken at a “slow” rate. Speaking rate was not measured quantitatively. The researcher spoke “fast” sentences as quickly as possible without sacrificing clarity and intelligibility. The researcher spoke “slow” sentences at a slower rate than normal speech. To control for variation in pronunciation of the made-up word caused by speaking rate, the production of the made-up word in each fast sentence was spliced into the corresponding slow sentence using the sound processing software Praat (“corresponding sentences” are further explained in and below Table 1 on the following page). This differs from Magloughlin’s (2018) methodology which spliced the initial consonant into the word of interest. We spliced entire words into sentences.

There were four variables of interest: speaking rate, word location in the sentence, number of syllables in the made-up word, and the vowel that was included in the made-up word. As discussed above, there were two options for speaking rate: fast vs. slow. For word location in a sentence, the word of interest could be the first word in the sentence or the last word in the sentence. There were two options for the number of syllables as well: one vs. two. The one-syllable (non-

filler) words always began with /tɹ/ followed by one of the vowels. The two-syllable (non-filler) words always began with a vowel followed by /tɹ/ and then were followed by the same initial vowel again. We included three vowels in this experiment: [ɑ], [ɪ], [ʌ]. These vowels were chosen because they are some of the English vowels that are most resistant to vowel reduction.

A Latin square approach was used to balance the number of sentences falling into each category. Table 1 below shows how the four groups were divided over the 16 sentences containing made-up words of interest. Each row represents one sentence. Because three vowels were used but there were four groups to fall into ([fast, initial, two], [fast, initial, one], [slow, initial, two], [slow, initial, one]), they could not be evenly distributed across the groups. Instead, [ɑ] was used in 8 sentences, [ʌ] in four sentences, and [ɪ] in four sentences.

Speaking rate	Word location	Syllable Count	Vowel	Speaking rate	Word location	Syllable Count	Vowel
fast	initial	two	[ʌ]	slow	initial	two	[ʌ]
fast	initial	two	[ɪ]	slow	initial	two	[ɪ]
fast	initial	one	[ʌ]	slow	initial	one	[ʌ]
fast	initial	one	[ɑ]	slow	initial	one	[ɑ]
fast	final	two	[ɑ]	slow	final	two	[ɑ]
fast	final	two	[ʌ]	slow	final	two	[ʌ]
fast	final	one	[ɑ]	slow	final	one	[ɑ]
fast	final	one	[ɪ]	slow	final	one	[ɪ]

Table 1: Independent Variable Values Across 16 Sentences of Interest

This distribution created pairs of sentences, according to sentence speed, as seen across the rows in Table 1. One sentence in the pair would be a fast sentence, and the other would be a slow

sentence. The sentences in the pair would match in terms of word location, syllable count, and vowel. Because the fast token was spliced into the slow sentences, for each fast vs slow sentence pairing, there is only one unique production of the word of interest. Thus, they match in that exact token as well. However, the other words in the sentence would be changed to better hide the obviousness of the pairing from the participants. For example, one pair is "She bought [ʌtrʌ]" and "I really like [ʌtrʌ]" where the first sentence was spoken slowly and the second quickly. Other than the speaking rate and English words, the sentences are the same with regard to the independent variables of interest in this experiment.

Participants

Participants were recruited using the Ohio State University Department of Linguistics Linguistics Outside the Classroom (LOC) participant pool. This experiment was listed on the LOC website where students can look for experiments to participate in. Students signed themselves up for experiments of their choosing and received course credit for experiment participation.

41 undergraduate OSU students participated in the study. Each filled out an anonymous questionnaire collecting information about age, gender, language knowledge, and residential history. The data from two participants were not included in the current study because of technical issues that resulted in missing data. Two other participants' data were also excluded because those participants had an audio or visual processing disorder. Data from 37 participants were included in the final analysis. Most ages ranged from 18-22, with individual participants aged 24, 25, 29, and 30. There were 27 women and 10 men. 17 students began learning a non-English language before the age of 5. All were native English speakers except for one student who began learning English at age 4. Multilingual speakers were not excluded for the sake of retaining a reasonable sample size.

Results

We analyzed the effects of the four independent variables (speaking rate, word location, syllable count, and vowel) on spelling choice. We compared the spelling choice responses using a generalized linear model, testing for main effects of the independent variables. Speaking rate did not have a statistically significant effect on spelling choice ($p > 0.05$). The figure below shows the rate at which participants chose the "ch" spelling option according to speaking rate, including a standard error range.

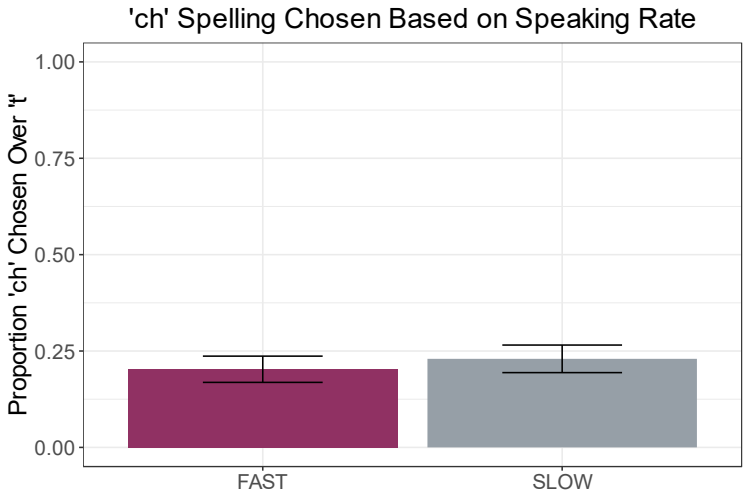

Figure 1. Effect of Speaking Rate on "ch" vs "t" Spelling Choice.

The vowel [ʌ] was found by the generalized linear model to have a statistically significant effect on spelling choice when compared to the vowel [ɑ] ($p < 0.05$). The vowel [ɪ] did not have a statistically significant effect on spelling choice when compared to the vowel [ɑ] ($p > 0.05$).

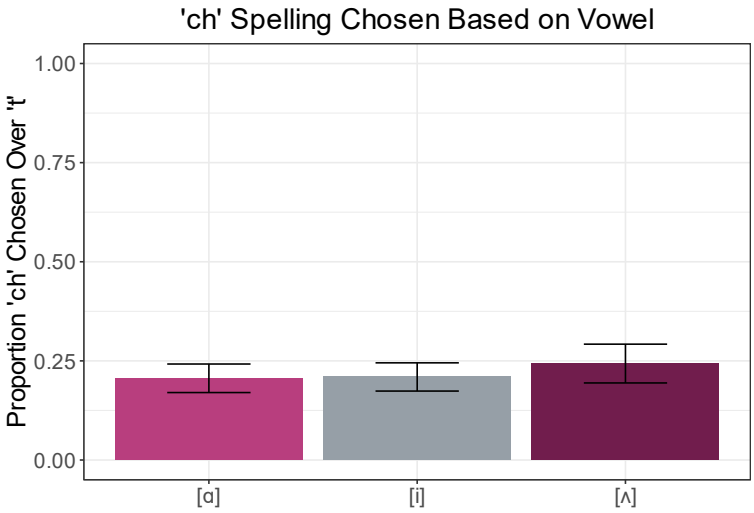

Figure 2. Effect of Vowel on "ch" vs "t" Spelling Choice.

However, both word location and syllable count produced statistically significant effects ($p < 0.05$). When the made-up word of interest was the first word in the sentence (e.g., "[ʌtrʌ] fell off the table"), participants were more likely to choose to spell the word with a "ch" than if the word was the last word in the sentence (e.g., "She bought [ʌtrʌ]"). This is also illustrated in Figure 3 below.

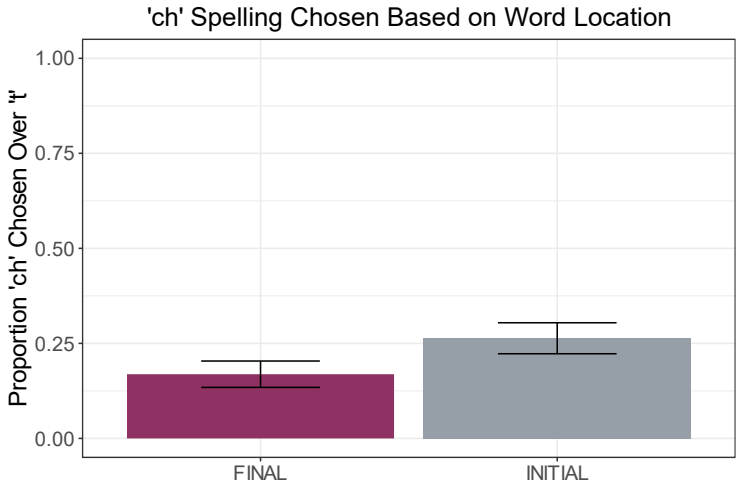

Figure 3. Effect of Word Location on "ch" vs "t" Spelling Choice.

When the made-up word of interest had two syllables in it (e.g., [ɑtrɑ]), participants were more likely to choose to spell the word with a “ch” than if the word had one syllable in it (e.g., [trɑ]).

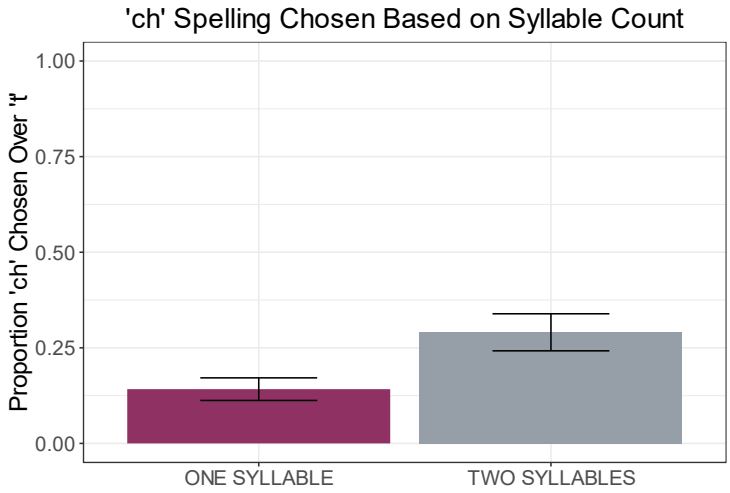

Figure 4. Effect of Syllable Count on “ch” vs “t” Spelling Choice.

Corpus Study

This experiment used a speech corpus to explore the effects of speaking rate, age, and gender on the affrication rate of /t/. Speaking rate and age are both applicable to the phonologization questions. Gender was included as a variable to explore any other possible interesting effects.

Buckeye Corpus of Conversational Speech

All data used for this corpus study were gathered from the Buckeye Corpus of Conversational Speech (Pitt et al., 2007). This corpus was constructed in order to create a database of spontaneously spoken words, lexically and phonetically transcribed and time-aligned. Work on this corpus began in 1999 when forty talkers were recruited to participate in 30-60 minute interviews with researchers from The Ohio State University in Columbus, Ohio. 300,000~ words were collected. All participants were native to the Central Ohio area and were middle-class

Caucasians. Participants were categorized based on two variables: gender (20 female and 20 male) and age (20 talkers under 30 years old and 20 talkers over 40 years old). Talkers under the age of 30 were born after 1970, and talkers over the age of 40 were born before 1960. During these sociolinguistic interviews, participants were asked to discuss topics like changes to Columbus over the years, sports, and raising children.

Trained transcribers first orthographically transcribed (i.e., wrote the text of) each interview by hand. The ESPS Aligner software was then used to automatically align dictionary pronunciations of the transcribed words with the content of the sound files. These alignments and phonetic transcriptions were then manually edited and improved by trained human labelers. The human labelers used spectrograms and waveforms of the sound files to aid their labeling. The end result was a set of files for each participant containing the phonetic transcription for every word they spoke, the dictionary pronunciation of that word, the orthographic transcription of that word, and the timestamp of when that word was spoken. This timestamp marks the end of the word. Separate files also contain timestamp information on the phone level, where the time for each individual phone in a word is recorded. For example, for the word 'play', timestamps are recorded for the phones [p], [l], and [ei] individually.

In the phonetic transcription protocol, corpus annotators were specifically called to pay attention to the possibility of palatalization (here, referred to as "affrication") of /t/ before /ɹ/, among other phenomena such as glottalization and nasalization. If /t/ was palatalized before /ɹ/, the /t/ was phonetically transcribed as "ch".

Methodology

A subset of these data was used for the purposes of this study. For all forty participants, only words containing "tr" in their orthographic spelling were included. This included words like

'traffic', 'electric', and 'strong'. Because affrication was attended to in the phonetic transcriptions, we were able to identify which productions included an affricated /t/. All "tr" containing words were categorized based on whether the /t/ had been affricated or not. To identify these, the phonetic transcription was used. If the phonetic transcription included "ch r", the /t/ was considered to be affricated. Otherwise, the word was not.

We calculated each participant's individual affrication rate. This was done on an individual basis because it relies on how many words each participant produced containing "tr". For example, a participant that produced 40 words containing "tr" and affricated 30 of them would have an affrication rate of 75%, but a participant that produced 50 words containing "tr" and affricated 30 of them would have an affrication rate of 60%.

Speaking Rate

In order to measure the effect of speaking rate on affrication rate, we have to first calculate the speaking rate for each individual word. This information is not directly included in the corpus. Measurement of speaking rate has been a point of debate for decades (Carroll, 1966). There are numerous approaches and complications (Mirghafori et al., 1996; Wang & Narayanan, 2007). For example, should long silences in the middle of a sentence be removed? One subset of approaches is counting the number of units per second, using units such as phones, syllables, words, etc. (Wang & Narayanan, 2007). These measurements require (phonetic) transcriptions of the speech being measured. However, these approaches may not be suitable for conversational speech as, due to their spontaneous nature, phones may be abnormally produced or even dropped (Morgan et al., 1997). More recent work has focused on new forms of speaking rate measurement using the speech signal itself rather than on transcriptions (Wang & Narayanan, 2007; De Jong & Wempe, 2007). Unfortunately, this data is not available in the corpus, so an estimation was required.

We chose phones per second as the speaking rate proxy because a better transcription-based measurement for conversational speech has not been identified. Additionally, it is more granular than words per second, and the information needed to calculate it was readily available. First, the word's timestamp was identified, marking the end time of the word. Then one second was subtracted from the marked time. The number of phones occurring within this one second range of time during the participant's conversation was counted. Speaking rate ranged from 2-24 phones per second. The average speaking rate was 13.05 phones per second.

Results

We analyzed the effects of the three independent variables (gender, age, and speaking rate) on affrication rate. We compared the affrication rates using a generalized linear model, testing for main effects of these independent variables. Gender did not have a statistically significant effect on affrication rate ($p > 0.05$). The figure below shows the rate at which participants produced an affricated /t/ according to their gender, including a standard error range.

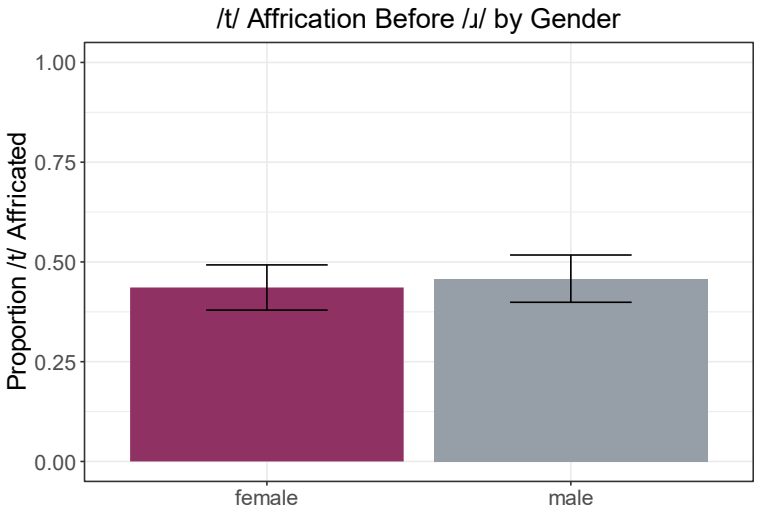

Figure 5. Effect of Gender on /t/ Affrication Rate.

Age did, however, produce a statistically significant effect on /t/ affrication ($p < 0.05$). Participants under the age of 30 had a higher rate of affrication (~52%) than the group of participants over the age of 40 (~38%).

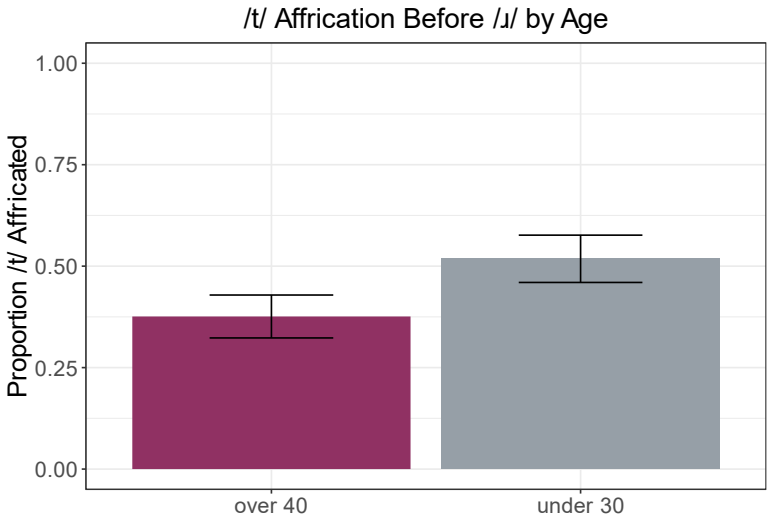

Figure 6. Effect of Age on /t/ Affrication Rate.

In the analysis of speaking rate effects, speaking rate is a continuous independent variable, and affrication is a binary dependent variable. Using the generalized linear model, speaking rate was found to have a statistically significant effect on whether the /t/ was affricated or not ($p < 0.05$). Higher speaking rates were more closely associated with lower rates of affrication. The average speaking rate around words containing an affricated /t/ was 12.75 phones/second, and the average speaking rate around words containing a non-affricated /t/ was 13.33 phones/second. Figure 7 shows affrication rate by speaking rate. There is a distinct downward trend in affrication as the speaking rate increases. Some speaking rates at the extreme ends of the range have affrication rates of 0% because there were only a few observed words at those speeds, and none were affricated.

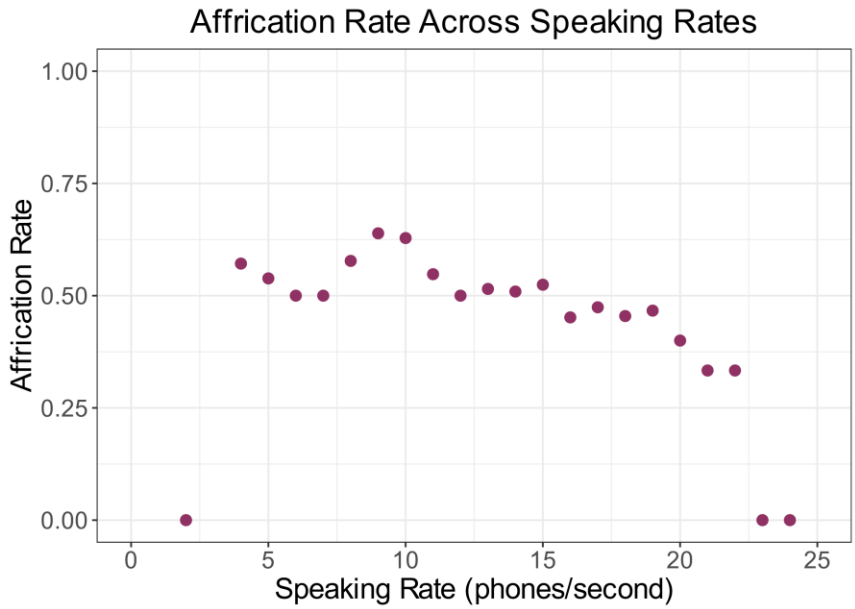

Figure 7. Effect of Speaking Rate on Affrication Rate.

Discussion

In the perception experiment, the rate of "ch" being the chosen spelling was always low, around or below 25%. This low rate was expected as it is almost identical to Magloughlin's (2018) third initial consonant type (/t/ spliced out of a recording where it appeared before /ɹ/) in the /ɹ/-following environment, which was overwhelmingly spelled using "t" in that study. The question was whether any of the four phonetic independent variables would influence this low rate. No statistically significant effect of speaking rate on spelling choice was found. Because perception of the affrication phenomenon appears to be resistant to phonetic influence that comes with a varying speaking rate, this result would support the conclusion that this phenomenon, /t/ → [tʃ] before /ɹ/, has been phonologized. However, statistically significant effects of word location and syllable count on spelling choice were observed, which directly contradicts the speaking rate finding. Words located at the beginning of the sentence resulted in more "ch" spellings, and words containing two syllables resulted in more "ch" spellings. The significant effects of vowel choice,

where [ʌ] resulted in more "ch" spellings than [ɑ], were less straightforward as the effect appears to depend on the specific vowel.

Because the connections of word location, syllable count, and vowel to articulation is less direct, no clear explanations for these effects were found in the literature. However, hypotheses are offered here. Regarding word location, it is possible that words produced at the beginnings of sentences are produced more quickly. Regarding syllable count, it is also possible that the /t/ segment in a ˈV.tɪV (stress on first syllable) word may be produced more quickly if in an unstressed position (as it was in the productions used for this experiment). It was reasoned that a shorter /t/ segment might result in more "t" spellings because [t] is shorter in duration than [tʃ]. To explore this, the /t/ sequences of each constructed /tɪ/ word were measured. Durations ranged from 112 ms – 139 ms. No significant pattern was found between the word initial vs. word final groups, so that effect is still left unexplained. However, for every pair of words that differed only by syllable count (e.g., both words were produced word-finally but one had one syllable and the other had two), the utterance with two syllables contained the shorter duration /t/ segment (and more "ch" spellings). This effect is in the opposite direction as the above prediction, with shorter /t/ segments resulting in more "ch" spellings. This phonetic difference in /t/ duration may have affected the perception of /t/ as a function of syllable count (or, perhaps more accurately, stress position). This would be evidence against phonologization. The differences in spelling choice for [ʌ] vs. [ɑ] vowel words may also be explained by duration. The words containing the vowel [ʌ] had the shortest overall /t/ durations, perhaps because [ʌ] is a central vowel from/to which it is easy for the tongue to move quickly. This matches the pattern of the syllable count effect, where the shorter duration /t/ segments resulted in more "ch" spellings. Overall, the perception study provides mixed results both in support of the phonologization conclusion and against it.

The corpus study reveals significant effects from age and speaking rate on affrication rate and no effects from gender. The effects from age, with younger talkers producing affrication at a higher rate than the older talkers, support Magloughlin's corpus findings which suggest that this phenomenon has been undergoing gradual change.

The effects found from speaking rate were rather unexpected. A finding revealing no effect from speaking rate would have supported the phonologization conclusion. A finding revealing higher affrication rates at higher speaking rates would have been evidence against the phonologization conclusion, supporting a phonetic effect of more coarticulation at faster speeds. However, a third option was found: lower affrication rates at higher speaking rates. No possible phonetic motivation for the direction of this effect has been identified, assuming that this effect is an accurate representation of the talkers' productions. Because it is still an effect on production from a phonetic factor, this is evidence against phonologization. However, the unexpected direction of the results may have been influenced by the data measurement techniques. The usage of phones per second as a proxy for speaking rate may not have been as accurate as it would need to be to detect true effects from speaking rate. As previously discussed, phones per second is not the optimal choice for speaking rate measurement for spontaneous, conversational speech. It was, however, the most appropriate proxy that was accessible through the corpus data. As it stands, the results from the corpus study provide evidence against phonologization.

Conclusion

As noted by Magloughlin (2018), /t/ affrication before /ɹ/ has long been a topic that is taken for granted in the literature and discussed in blogposts (LarsH, 2001; Read, 1971; Smith, 2013; user3604, 2013). As such, it has not often been the subject of rigorous scientific analysis. This paper, following the footsteps of Magloughlin (2018), aimed to continue filling that gap. Their

2018 work, using both a corpus study and perception study, provided evidence leaning in support of the idea that this phenomenon has been phonologized. However, the results from the current study complicate this conclusion.

While the lack of a significant effect of speaking rate in the perception study supports the notion of phonologization of the /t/ affrication before /ɹ/ phenomenon, significant effects of word location, syllable count, and vowel choice provide evidence against this conclusion. It may be possible that different phonetic processes, such as speaking rate vs. stress-induced differences from syllable count, affect the affrication of /t/ to different degrees, which would still result in a rejection of the phonologization conclusion. However, speaking rate was the phonetic factor of interest in the corpus study, and its results also argue against phonologization.

Despite efforts to provide a clear answer to the phonologization question, the status of the /t/ affrication before /ɹ/ phenomenon is still not entirely clear. The results from the two experiments provide conflicting evidence regarding the phonologization question, with some support for phonologization from the perception experiment (from speaking rate) and support for rejection of phonologization from both the perception study and the corpus study. Evidence may be leaning against the phonologization of this phenomenon. This stands in opposition to Magloughlin's 2018 findings. Thus, the question of whether the affrication phenomenon has undergone phonologization remains open, warranting further investigation with methodological refinement.

References

Agwuele, A., Sussman, H. M., & Lindblom, B. (2009). The effect of speaking rate on consonant vowel coarticulation. *Phonetica, 65*(4), 194-209.

Anttila, A. (2018). Variation in phonology. In *Oxford Research Encyclopedia of Linguistics*.

Bermúdez-Otero, R., & Trousdale, G. (2012). Expanded version of section 2 of: 'Cycles and continua: on unidirectionality and gradualness in language change'. In Terttu Nevalainen and Elizabeth Closs Traugott (eds), The Oxford handbook of the history of English, 691-720. New York: Oxford University Press.

Boomershine, A., Hall, K. C., Hume, E., & Johnson, K. (2008). The impact of allophony versus contrast on speech perception. *Contrast in phonology*, 143-172.

Carroll, J. B. (1966). PROBLEMS OF MEASURING SPEECH RATE.

Celata, C. (2007). Partial phonological contrasts in native and non-native speech perception.

De Jong, N. H., & Wempe, T. (2007). Automatic measurement of speech rate in spoken Dutch. *ACLC Work Pap, 2*, 49-58.

Dilley, L. C., & Pitt, M. A. (2010). Altering Context Speech Rate Can Cause Words to Appear or Disappear. *Psychological Science, 21*(11), 1664-1670. https://doi.org/10.1177/0956797610384743.

Gay, T., Ushijima, T., Hiroset, H., & Cooper, F. S. (1974). Effect of speaking rate on labial consonant-vowel articulation. *Journal of Phonetics, 2*(1), 47-63.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r." *Neuropsychologia, 9*, 317-323.

Hume, E., & Johnson, K. (2003). The impact of partial phonological contrast on speech perception.

In *Proceedings of the 15th international congress of phonetic sciences* (pp. 2385-2388). Barcelona: Universitat Autónoma de Barcelona.

Hyman, L. M. (2013). Enlarging the scope of phonologization. *Origins of sound change: Approaches to phonologization*, 3–28.

Kilbourn-Ceron, O., & Sonderegger, M. (2018). Boundary phenomena and variability in Japanese high vowel devoicing. *Natural Language & Linguistic Theory, 36*, 175-217.

Kiparsky, P. (2013). Phonologization.

[LarsH]. (2011, December 1). *Why are /t/ and /d/ sometimes affricated before /ɹ/ in English* [Online forum post]. Linguistics StackExchange. https://linguistics.stackexchange.com/questions/1104/why-are-t-and-d-sometimes-affricated-before-%c9%b9-in-english

MacKain, K., Best, C., & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics, 2*(4), 369-390.

Magloughlin, L. (2018). */tɹ/ and /dɹ/ in North American English: Phonologization of a Coarticulatory Effect.* Ph.D. thesis, University of Ottawa, Ottawa, ON, Canada.

Mirghafori, N., Fosler, E., & Morgan, N. (1996, May). Towards robustness to fast speech in ASR. In *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings* (Vol. 1, pp. 335-338). IEEE.

Morgan, N., Fosler-Lussier, E., & Mirghafori, N. (1997, September). Speech recognition using on-line estimation of speaking rate. In *Eurospeech* (Vol. 97, pp. 2079-2082).

Phillips, J. (2018). Temporal dynamics of /s/-retraction in American English.

Pitt, M.A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E. & Fosler-Lussier, E.

(2007). Buckeye Corpus of Conversational Speech (2nd release) [www.buckeyecorpus.osu.edu] Columbus, OH: Department of Psychology, Ohio State University.

Read, C. (1971). Pre-school children's knowledge of English phonology. *Harvard educational review, 41*(1), 1-34.

Reinisch, E., Jesse, A., & McQueen, J. M. (2011). Speaking rate affects the perception of duration as a suprasegmental lexical-stress cue. *Language and Speech, 54*(2), 147-165.

Scudieri, K. (2012). *On The Perception of /s/ and /ʃ/: Considering the Effects of Phonotactics.* The Ohio State University Department of Linguistics Honors Theses, Arts and Sciences Undergraduate Research Theses and Honors Research Theses.

Smith, B. J. (2013). *The interaction of speech perception and production in laboratory sound change*. The Ohio State University.

Summerfield, Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of experimental psychology. Human perception and performance, 7*(5), 1074–1095. https://doi.org/10.1037//0096-1523.7.5.1074

[user3604]. (2013, May 18). *A question about /t/ + fricative /r/ vs /tr/ affrication* [Online forum post]. Linguistics StackExchange. https://linguistics.stackexchange.com/questions/7201/a-question-about-t-fricative-r-vs-tr-affrication

Wang, D., & Narayanan, S. S. (2007). Robust speech rate estimation for spontaneous speech. *IEEE Transactions on Audio, Speech, and Language Processing, 15*(8), 2190-2201.

Yu, A. C. (2021). Toward an individual-difference perspective on phonologization. *Glossa: a journal of general linguistics, 6*(1).