

# Filozofická fakulta Univerzita Karlova

Bakalárska práca



Praha, 2024

Dorota Waczlavová

Filozofická fakulta

Univerzita Karlova

Ústav anglického jazyka a didaktiky

Bakalárska práca

Dorota Waczlavová

**Linking in British and American political debates**

Viazanie v britských a amerických politických debatách

Praha 2024

doc. Mgr. Radek Skarnitzl, Ph.D

## **Acknowledgments**

I want to thank my supervisor, doc. Mgr. Radek Skarnitzl, Ph.D., for his guidance and knowledge, which helped me write this BA thesis, for his feedback, instructions, and willingness to help whenever needed.

Prehlasujem, že som bakalársku prácu vypracovala samostatne, že som riadne citovala všetky použité pramene a literatúru, a že práca nebola využitá v rámci iného vysokoškolského štúdia, či k získaniu iného alebo rovnakého titulu. Pri úprave textu som použila umelú inteligenciu.

I declare that the following BA thesis is my own work, for which I used only the secondary literature that is listed in the resources. This thesis was not used as a part of any other university study, nor was it used to gain a different university degree. I have used artificial intelligence to adjust the text.

## **Abstract**

The aim of this thesis was to investigate linking in spontaneous speech in native English speakers. The main questions that we focused on were: what are the general and individual tendencies of speakers of English for linking, and what are the differences in these tendencies in speakers of British English compared to speakers of American English. We examined 16 native speakers; 8 of them were speakers of American English, and 8 of them were speakers of British English. Both of these groups were further divided into 4 male speakers and 4 female speakers. We analyzed these recordings using the computer program Praat. Firstly, we adjusted the boundaries of beginnings and endings of words. Then, we focused on the analysis of vowel sounds, looking closely at whether they are linked or glottalized. We extracted data from the analyzed recordings and put them into graphs so the results could be seen clearly. We found out that American and British English speakers favored consonant-to-vowel linking more than vowel-to-vowel linking. We also learned how vital speech rate is when influencing linking. British speakers tended to have a faster speech rate, resulting in frequent linking. On the other hand, American speakers had a slower speech rate, resulting in fewer instances of linking.

**Key words:** English, American, British, linking, connected speech, speech rate, spontaneous speech

## **Abstrakt**

Hlavným cieľom tejto práce bolo preskúmať viazania v spontánnej reči rodených hovoriacich. Hlavné otázky, na ktoré sme sa zamerali boli: aké sú obecné a individuálne tendencie anglicky hovoriacich k viazaniu, a aké sú rozdiely v týchto tendenciách u hovoriacich britskej angličtiny v porovnaní s hovoriacimi americkej angličtiny. Skúmali sme 16 rodených hovoriacich; 8 z nich boli hovoriaci americkej angličtiny a 8 z nich boli hovoriaci britskej angličtiny. Obe tieto skupiny sme ďalej rozdelili na 4 mužov a 4 ženy. Tieto nahrávky sme analyzovali pomocou počítačového programu Praat. Najprv sme upravili začiatky a konce slov. Potom sme sa sústredili na analýzu samohlások, pričom sme pozorne sledovali, či sú zviazané alebo glotalizované. Z analyzovaných nahrávok sme vytiahli dáta, ktoré sme dali do grafov, aby sme výsledky mohli prehľadne zobrazit'. Zistili sme, že americkí a britskí hovoriaci preferujú viazanie samohlásky so spoluhláskou viac ako viazanie samohlásky so samohláskou. Taktiež sme sa dozvedeli aká dôležitá je rýchlosť reči a jej vplyv na viazanie. Britskí hovoriaci mali tendenciu rýchlejšieho tempa reči, čo viedlo k častejšiemu viazaniu. Na druhej strane, americkí hovoriaci mali pomalšie tempo reči, čo viedlo k menej častým prípadom viazania.

**Kľúčové slová:** Angličtina, Americká, Britská, viazanie, spojená reč, tempo reči, spontánna reč

## **Table of Contents**

1. Introduction .....	8
2. Theoretical Background .....	9
2.1 Spontaneous Speech .....	9
2.1.1 Variability .....	12
2.1.2 Predictability .....	13
2.1.3 Prosody .....	14
2.1.4 Style-shifting .....	16
2.2 Connected Speech .....	17
2.2.1 Assimilation .....	17
2.2.2 Elision .....	22
2.2.3 Linking .....	23
3. Materials and methodology .....	28
4. Results and discussion .....	33
5. General Discussion and conclusion .....	43
6. References .....	45
7. Resumé .....	46

## List of Figures

<b>Figure 1.</b> Occurrence of glottalization in the American data. ....	29
<b>Figure 2.</b> Occurrence of glottalization in the American data. ....	30
<b>Figure 3.</b> Occurrence of linking in the American data. ....	31
<b>Figure 4.</b> Bar chart showing the overall number of linking vs glottalization for speakers of American and British English. ....	33
<b>Figure 5.</b> Percentage bar chart demonstrating the distribution of linking and glottalization for speakers American and British English. ....	33
<b>Figure 6.</b> Bar chart showing the overall number of linking vs glottalization for speakers of American and British English per word type. ....	35
<b>Figure 7.</b> Percentage bar chart demonstrating the distribution of linking and glottalization per word type. ....	35
<b>Figure 8.</b> Bar chart demonstrating the overall number of linking vs glottalization for individual speakers. ....	37
<b>Figure 9.</b> Percentage bar chart demonstrating the distribution of linking and glottalization for all speakers. ....	37
<b>Figure 10.</b> Percentage bar chart demonstrating the distribution of linking and glottalization per word type for individual speakers. ....	39
<b>Figure 11.</b> Bar chart showing the overall number of linking vs glottalization for speakers of American and British English specifically focused on consonants and vowels. ....	40
<b>Figure 12.</b> Percentage bar chart showing the overall number of linking vs glottalization for speakers of American and British English specifically focused on consonants and vowels. ....	41



## **1. Introduction**

This study is aimed to research linking in spontaneous speech to find out whether the existing research can also be applied to this type of speech. We wanted to examine how native speakers of British and American English link in spontaneous speech and to investigate whether the textbook rules for linking are present. Many studies touch on the process of linking with materials recorded in laboratory settings only, so we believe researching spontaneous speech may help us understand linking more clearly.

Materials used in this research consist of 16 recordings; 8 of them are of American English speakers, and 8 of them are of British English speakers. We looked at the aspect of linking in each of the recordings. It was interesting to see how spontaneous speech works with linking, and to see the differences between "textbook" linking and linking in "real" life. It was also interesting to look at the differences in linking between the two types of English, American, and British. We compared the amount of glottalization and linking present in the recordings, focused on semantics and word type, and other aspects present in phonetic context.

There are many different studies, whether recent or older, that focus on the process of linking. Unfortunately, there are not as many of them that investigate linking in spontaneous speech or in native speakers. We thought it would be productive to conduct this research to find out how linking acts in spontaneous speech and what the general and individual tendencies of native speakers are.

The theoretical part that follows this introduction will give you an idea of what is fluent speech, spontaneous speech and what are the aspects and rules of connected speech. It will introduce the topic of linking, what type of linking we can arrive at when analyzing speech and what are the tendencies of linking.

## **2. Theoretical Background**

### **2.1 Spontaneous Speech**

When we analyze language, we can examine it from various points of view. We can examine the morphologic structure, the syntactic structure, the phonetic structure, etc. The topic of this thesis is linking. To arrive at this aspect of connected speech processes, I will take a broader look at language in more general terms, focusing firstly on fluency, spontaneity, and connected speech processes.

Fluency of language does not only relate to spoken language or our perception of it, as one might think. Fluency is a mental process; it is our understanding of certain things, and it is how we interpret information. If the task is easy and fluent, it can be "solved" quicker. As Oppenheimer states, "fluency is defined as a subjective experience of ease or difficulty associated with a mental process. In other words, fluency is not the process itself but information about how efficient or easy that process feels" (Oppenheimer, 2008, p. 238).

The difficulty or ease with which we interpret information can be shaped in multiple ways. We can make the process of interpretation easier or harder depending on what we want to achieve. We can always manipulate fluency, e.g., when reading a text, the font of it can be smaller, so the difficulty of comprehending the information becomes challenging (Oppenheimer, 2008).

When defining spontaneous speech, we have to look at numerous characteristics that define it. We can look at spontaneous speech as containing numerous different speech styles. Spontaneous speech serves as an umbrella term for many speech styles that can be put under it. What is interesting to note is that the term "speech style" does not have a set classification. Some researchers may talk about connected speech and could mean two different things, "in one case this could mean carefully read sentences, while in another case this could mean a conversation between two speakers" (Tucker & Mukai, 2023, p. 3). As Tucker and Mukai

(2023) state, “we define speech style as a form of language produced on the basis of internal and external factors. These factors could depend on situation, formality, mood, individual choice, and environment” (Tucker & Mukai, 2023, p. 2, 3). According to this, spontaneous speech can be then considered as “conversational, connected, casual, fast, natural, and vernacular” (Tucker & Mukai, 2023, p. 3).

Defining spontaneous speech is difficult; one person can have different specifics than another, making it challenging to strictly define this term. Tucker and Mukai (2023) touch briefly on the issue of defining spontaneity and state that spontaneous speech is a speech that is not prepared. What is interesting to note is how they divide this in terms of speech styles. They talk about three parts: 1) careful to casual speech, 2) rate, and 3) reduced to unreduced (Tucker & Mukai, 2023). Careful speech can be defined as a type of speech where we are highly aware of our pronunciation and the overall process of our speech production. Casual speech, on the other hand, shows higher variability, articulation rate, and more frequent hesitations. The rate category relates to how fast our speech becomes in, e.g., conversations. The third category concerns reduced forms of words, and the choice of whether to use reduction depends primarily on external factors. All of these parts are included in spontaneous speech, and they can be but do not have to be combined, e.g., having a casual conversation with a friend but needing to talk slowly while using unreduced forms because of external factors. What Tucker and Mukai (2023) do is create a definition that encompasses every term that previous researchers have come up with and group it into one. "Spontaneous speech is speech produced by a speaker in an informal, dynamic, unrehearsed, casual manner" (Tucker & Mukai, 2023, p. 5). Spontaneous speech is not easily defined, and while the definition provided by Tucker and Mukai (2023) is helpful, it does not have to agree with some situations that other people could see as spontaneous or not spontaneous.

Spontaneous speech can be defined as the opposite of careful, laboratory, or read speech. Traditionally, language has been studied primarily in a laboratory environment. The study started with pseudo-words, or "nonsense" words, that were used to closely examine speech. Later, scientists added short sentences and isolated words. These were recorded under strict conditions in a laboratory setting. All the information that was gathered by these investigations has been essential in the study of speech but was limited only to speech that, as mentioned, was examined under conditions that could be produced only in phonetic laboratories. We can only sometimes apply these findings when it comes to spontaneous speech. Tucker and Mukai (2023) focus on the research done by William Labov (1972) and the importance of the study of spontaneous speech, which mentioned in his work is "to obtain the data most important for linguistic theory, we have to observe how people speak when they are not observed" (Labov, 1972, p. 113, cited by Tucker & Mukai, 2023, p. 2). Nowadays, there are more studies that would use materials that are natural and are recorded with the intention to catch spontaneous speech in its natural environment, moving away from the laboratory research, but unfortunately only a few that would focus on a specific phonetic feature of spontaneous speech.

Another essential aspect to note is the use of canonical forms of words in spontaneous speech. The term canonical form is used when talking about a form of a word that is used as its standard representation. The standard representation of a word, in this case, its pronunciation, can be found in a dictionary and can be described as a careful pronunciation of its form. Many researchers use the canonical form as the one from which all others deviate. This is productive mainly when searching for reduction or strengthening of one form, as the comparison to the canonical form shows the difference. The question is whether using the canonical form as the norm is beneficial in these studies. When looking at all of the definitions of spontaneous speech above, one has to note that in most cases, the canonical form never occurs. As Tucker and Mukai (2023) state, "if we, however focus not on reduction or strengthening but on the

variation of the forms produced, we will likely find that the most common form offers a production and processing advantages and that disadvantages are found for the less-common-reduced and extremely careful forms” (Tucker & Mukai, 2023, p. 6).

### **2.1.1 Variability**

Variability of speech plays a vital role in spontaneous speech. Spontaneous speech is inherently variable, concerning different variations of words, segments and syllables. Variability is, at its core, the deviation from the canonical form of words (Tucker & Mukai, 2023, p. 1). Their "dictionary" pronunciation is highly changeable, mainly when we talk about fluent, spontaneous speech. In comparison to careful read speech, where the pronunciation almost always includes the canonical form, spontaneous speech does not. Reduction is another aspect that concerns variability. Here, we talk about the acts of deletion, shortening, or incomplete pronunciation (Tucker & Mukai, 2023).

Tucker and Mukai (2023) focus on the phenomenon of variability of speech and discuss works by Johnson (2004) and Dilts (2013), where they both investigated the Buckeye Corpus with focus on the deviation from the canonical form. "Johnson (2004) found that over 25 percent of the words in a subset of the Buckeye Corpus are missing phones or segments based on comparisons to the canonical form" (Tucker & Mukai, 2023, p. 16). Dilts (2013), continuing in this research, found out that these deviations from the canonical form make over 32 percent of the Buckeye Corpus (Tucker & Mukai, 2023).

When we talk about variability, the question of consistency of speech arises. This also closely intertwines with the topic of context related to variability. According to the results provided by Tucker and Mukai (2023), the answers vary. When we consider careful read speech, and as mentioned above, it naturally has less variability. Variability is heavily language-dependent, sometimes language-independent. All of this shows us that variability is both consistent and inconsistent (Tucker & Mukai, 2023).

As mentioned above, variability closely connects to the topic of context. Tucker and Mukai (2023) talk about context as, “it includes the environment, the ambient noise, the interlocutor, social factors, and many other possible aspects of conversation” (Tucker & Mukai, 2023, p. 20). If we talk about the realization of variability, we also have to talk about the contextual factors that are involved in the process since many of them help us answer the questions we have about consistency. Multiple contextual conditions influence the realization of variability, although in this thesis, I will mention only two: predictability and prosody.

### **2.1.2 Predictability**

Predictability helps us predict spontaneous speech by predicting how much reduction will occur in a conversation if the duration of words is shorter due to predictability (Tucker & Mukai, 2023).

When it comes to predictability, some studies "seek to quantify the predictability of certain parts of the speech and use this quantification to make predictions about the variability in the speech signal. These studies focus on reduction and use context to identify consistent characteristics of phonetic variability” (Tucker & Mukai, 2023, p. 20). Such studies have been done on, for example, the Dutch language, where the focus of the research was on predictability in cases of informativity and acoustic characteristics of speech sounds. Tucker and Mukai (2023) discuss the research of van Son et al. (1998), where they “found a consistent relationship between informativity (measured as frequency) and the acoustic characteristics of speech sounds (duration and spectral properties)” (Tucker & Mukai, 2023, p. 21). The specific results included the findings “that the duration and spectral characteristics are reduced when occurring in highly predictable sequences and enhanced/strengthened in low predictable sequences” (Tucker & Mukai, 2023, p. 21).

Speech variability is also influenced on a morphological level by contextual predictability, and this has been shown in a study by Pluymackers et al. (2005b) done again on

the Dutch language, which Tucker and Mukai (2023) use as another example study. This particular study focused on affixes and “found that the frequency of the stem (eigen ‘own’ in ont-eigenen ‘to disown’) influences affix duration (ont-) in Dutch – with higher frequency stems predicting shorter affixes” (Tucker & Mukai, 2023, p. 21, 22).

The last study Tucker and Mukai (2023) discuss on the topic of speech predictability is research done by Hanique and Ernestus (2011). This study was done on the Dutch language again and focuses on reduction and word-final /t/ in past participles and how this "is predictable from the frequency of the preceding two words" (Tucker & Mukai, 2023, p. 22). The other finding of this study was "that the frequency of the word relative to its lemma frequency (the total frequency of all forms of a lemma, e.g., jump, jumps, jumped, etc.) is predictive of word-final /t/ reduction" (Tucker & Mukai, 2023, p. 22).

### **2.1.3 Prosody**

Prosody is a second contextual condition with which we can predict spontaneous speech. Similarly to predictability, prosody is interested in duration. The combination of redundancy and prosodic structure "predict durational variation in spontaneous speech" (Tucker & Mukai, 2023, p. 23). Tucker and Mukai (2023) refer to research done by Aylett and Turk (2014), where they claim that there exists “an inverse relationship between language redundancy and duration, a relationship between prosodic prominence and duration, and that much of the durational variance is accounted for by the redundancy and prosodic prominence measures” (Tucker & Mukai, 2023, p. 23).

What is important to mention is how we can categorize prosody, or speech rate, into three parts. Tucker and Mukai (2023) use the classification provided in the work of Crystal & House (1990). The first type of speech rate is “the production rate or speech rate calculated as the number of production units (e.g., word, syllables or phones) per unit time (e.g., duration of the word or utterance) including pauses" (Tucker & Mukai, 2023, p. 23). The second type is

“the articulation rate which is calculated as the number of production units (e.g., word, syllables or phones) per unit time (e.g., duration of the word or utterance) excluding pauses,” and the third type includes “the pause rate calculated as the number of pauses per duration of the utterance” (Tucker & Mukai, 2023, p. 23).

Tucker and Mukai (2023) use a number of different, interesting studies in their work to illustrate how speech rate works as a contextual factor in speech variability. I have decided to use three more to further explain how prosody is used for predicting spontaneous speech.

Works by Kohler (1996) and Trouvain et al. (2001), discussed in Tucker and Mukai (2023), used the German Kiel Corpus for Read and Spontaneous Speech for research on the articulation rate of spontaneous and read speech. They "found that spontaneous speech is produced with a faster articulation rate than read speech and that spontaneous speech contains greater variability" (Tucker & Mukai, 2023, p. 24). This study will be relevant mainly in the practical part of this thesis, where the articulation rate of spontaneous speech plays a vital role in the results.

They also found that even though spontaneous speech has a higher articulatory rate on average, “it contains a high number of slow utterances, possibly due to the large number of very short utterances” (Tucker & Mukai, 2023, p. 24). The results that come from this research also makes Tucker and Mukai (2022) “suspect that the increased variability in speech rate is likely due to the speaker using rate as a contextual cue to convey additional information about the speech and to provide additional clarity when necessary” (Tucker & Mukai, 2023, p. 25).

Tucker and Mukai (2023) also focus on prosodic boundary markers using research done by Blaauw (1994), which compared these markers in read and spontaneous speech. The first result of the study showed that “additional pauses, creating a full intonational boundary, occur more in spontaneous than in read speech” (Tucker & Mukai, 2023, p. 24). The second result displayed that "phrase-internal boundaries are typically realized without boundary-marking



pitch movements in spontaneous speech" (Tucker & Mukai, 2023, p. 24). Lastly, the results showed that "falling boundary tones are more frequent in read than in spontaneous speech" (Tucker & Mukai, 2023, p. 24).

Prosody is an important aspect of spontaneous speech and can help in its analysis. It provides information about the context of speech that can help in interpreting it. Tucker and Mukai (2022) discuss the research done by Mehta and Cutler (1988), which found that prosody greatly helps when distinguishing between two speech styles. They analyzed read and casual speech, where the prosodic features influenced listeners to react quicker in read speech rather than in spontaneous speech.

#### **2.1.4 Style shifting**

The last thing I would like to briefly touch upon in this part is style shifting. This idea is used in a significant amount of research that I will use in this thesis. Both Tucker and Mukai (2023), Hieke (1984), Alameen (2007) and Cruttenden (2014) discuss the importance of style shifting in spontaneous speech. I already mentioned the idea of speech styles, e.g. casual and read speech. When the speakers take part in a casual speech, the increase in reductions is immense; many times, the speaker is in a comfortable situation where he does not have to watch the exact pronunciation of words and overall phrasing of his speech. Alameen (2007) claims that, "when the speaker and the listener both belong to the same social group and share similar speech conventions, the comprehension load on the listeners will be reduced, allowing them to pay less attention to distinctive articulation" (Alameen, 2007, p. 12). Careful read speech, or careful speech in general is precisely the opposite. The speaker takes special care when choosing the correct form of a particular word or thinks carefully about the phrasing he intends to use.

## **2.2 Connected Speech**

I have introduced the idea of fluency and various aspects of spontaneous speech. Let me slightly narrow this topic and move to what connected speech processes are. Written and spoken language are in their essence very different. Written texts have their grammar, certain spellings of the words. The words may have a completely different pronunciation when compared to the written form in spoken language. When producing spoken language, especially natural speech in conversations, we notice that for our speech to sound natural, we need to make it flow. In natural speech, we often times omit endings or beginnings of words, such as in *this soap* [ðɪs səʊp] where we omit [s] to make it blend into the following sound. When we make speech connected, we inherently change the structure of the word. Changes of words in speech include additions, omissions, and even complete changes of sounds. In this way we can say that the words are undergoing changes and modifications. When we hear a natural conversation between two or more speakers, we notice the words glide into one another. We do not speak every word with a pause after it; we would inevitably sound like robots if we did. In this part of my work, I will dissect every process that creates connected speech and compare different views of scientists on this topic.

As I mentioned before, scientists may use different terms for the same processes. They either create different names for the processes of connected speech, or they create new categories altogether. The base of this introductory part is going to be the research of Roach (2009) which I am going to compare to the research of Alameen and Levis (2015) and that of Cruttenden (2014).

### **2.2.1 Assimilation**

Assimilation is a connected speech process that influences how words that are close to each other are pronounced. It is more likely to see assimilation occur in fast speech or natural conversation than in slow speech, where the speaker enunciates each word carefully. When we

are speaking, we do not produce each word in isolation. Assimilation affects mainly consonants; one phoneme changes into another because it is in close proximity to another phoneme of another word (Roach, 2009). Assimilation is divided into two other categories: regressive and progressive. When it comes to regressive assimilation, the phoneme of a preceding word becomes affected by the one phoneme of a following word, as, for example, *in bed* is pronounced [ɪm bed]. Progressive assimilation is the opposite, phoneme of the following sound becomes affected by the phoneme of the preceding sound. Example of progressive assimilation would be *in the* [ɪn ðə] where [ð] becomes [n] as in [ɪn̩nə].

When we talk about assimilation, I have to add that since assimilation affects consonants, we also talk about assimilation in terms of place, manner, and voicing.

The example of assimilation of place can be seen in my previous example [ɪm bed]. We can see that *in* is ending with [n], is a consonant which has an alveolar place of articulation, *bed* beginning with [b] is a bilabial consonant which changes [n] into [m], changing it from alveolar to bilabial in consequence. We can also see the effects of regressive assimilation of place in other sounds, such as [d] to [b] before [b,m,p], e.g. *good boy* [gʊb<sup>7</sup> bɔɪ]. [d] also assimilates into [g] when following [k,g], e.g. *good cake* [gʊg<sup>7</sup> keɪk]. [t] becomes [p] before [b,m,p] in e.g. *that prune, that ball* [ðæp<sup>7</sup> pru:n, ðæp<sup>7</sup> bɔ:l]. When we have an alveolar consonant followed by a velar stop, alveolar becomes velar as in *in Greece* [ɪŋ gri:s]. The transformation of an alveolar consonant to a dental plosive when dental consonant follows as can be seen in an example *that thought* [ðæt̪ θɔ:t]. Lastly there is the change of alveolar fricative followed by postalveolar which makes postalveolar fricative. This includes sounds s and z, which transform into [ʃ] and [ʒ] when followed by them, as in *this shop* [ðɪʃ ʃɒp]. Roach (2009) states that, “Assimilation of place is only noticeable in this regressive assimilation of alveolar consonants” (Roach, 2009, p.111).

On the other hand, differences in the assimilation of manner are not so significant as in assimilation of place. Assimilation of manner “is only found in the most rapid and casual speech; generally speaking, the tendency is again for regressive assimilation, and the change in manner is most likely to be towards an "easier" consonant - one which makes less obstruction to the airflow" (Roach, 2009, p. 111, 112). Here we yet again divide it into progressive and regressive assimilation. Regressive assimilation of manner, where plosive becomes a fricative or a nasal, can be seen, for example *on, good night* becoming [gʊn naɪt]. We can find progressive assimilation of manner in words beginning with [ð]. Nasal or plosive is followed by [ð] and becomes identical to the fricative but has a more dental place of articulation (Roach, 2009), for example, *in the* [ɪn̪ðə]. The last change of consonants can be seen in voicing, but again these can be found only in certain instances. Cruttenden (2014) also mentions in his work assimilation of manner but calls it a bit differently than Roach (2009). Focusing mainly on nasality and labialisation, phonemic assimilation that includes nasality affects alveolars, mainly when they are connected to not, or the contracted form *n't*. Initial voiced consonant, which is usually a plosive, transforms into a nasal. (Cruttenden, 2014) The most typical changes are from [d] to [n], [v] to [m], etc. Here we can see it on the example *She wouldn't dance* where [d] becomes [n], [ʃi: 'wʊn(t) dɑ:ns]. It is also important to mention that when it comes to labialization, this does not make phonemic changes, although we can see that there is a difference in lip rounding, resulting in different lengths in sound, e.g. [ɒ] and [ɑ:], but on this I will briefly touch on later.

When talking about voicing, we can have a word with a final consonant that is lenis, in other words voiced, and a word with an initial consonant that is fortis. In this case, the lenis consonant is voiceless, and we can see this phenomenon in the example such as *bed time*, where *bed* is pronounced with [d], but transforms into [d̥] as in [bed̥ taɪm], and as it is devoiced, it becomes almost a [t] in pronunciation. This is not as noticeable as “initial and final lenis

consonants usually have little or no voicing anyway; these devoiced consonants do not shorten preceding vowels as true fortis consonants do” (Roach, 2009, p. 112).

In the assimilation of voicing, only regressive assimilation is possible, which here has only one type. I have been talking mainly about assimilation regarding word boundaries. There is also assimilation that relates to morpheme boundaries. Progressive assimilation of voice, which forms the plural of 3<sup>rd</sup> person singular, suffixes -s and -z, has become fixed. “When a verb carries a third person singular '-s' suffix, or a noun carries an '-s' plural suffix or an '-s' possessive suffix, that suffix will be pronounced as s if the preceding consonant is fortis ("voiceless") and as z if the preceding consonant is lenis ("voiced")” (Roach, 2009, p.112). This “rule” can be seen on examples such as *waits*, *knocks* [weɪts, nɒks] and *hugs* and *trains* [hʌgz, treɪnz]. It is interesting to note that voiceless consonants that are in the final position do not assimilate to voiced, e.g., *black door* is not pronounced as [blag dɔː]; this phenomenon can be observed mainly in foreign speakers.

There is not only assimilation across word boundaries; similar changes can also be seen across morpheme boundaries. "If in a syllable-final consonant cluster, a nasal consonant precedes a plosive or a fricative in the same morpheme, then the place of articulation of the nasal is always determined by the place of articulation of the other consonant” (Roach, 2009, p.112), and this can be seen on examples *hump*, *dump*.

After dissecting how consonants change in certain situations, Cruttenden (2014) has a similar type of division as Roach (2009), but he also includes coarticulation that occurs in language. These allophonic variations, change of place and manner of articulation, and difference in voicing, Cruttenden (2014) also includes lip position and position of the soft palate. "Lip position is under the influence of adjacent vowels or semi-vowels" (Cruttenden, 2014, p.309). If we look at these changes within words, we find that, for example, [p] is lip-rounded in loop, but lip-spread in peak. The same can be seen, for example, in [ʃ], which is lip-

spread in sheep and lip-rounded in shook. Lip position changes can be seen not only within words but also at word boundaries. These influence for example [t] *that one*, [k] *thick one*, [s] *shall we*, and they are labialised in some form. The last part that Cruttenden (2014) mentions in allophonic variations is the position of the soft palate. "Nasal resonance - resulting particularly from regressive but also from progressive lowering of the soft palate in the vicinity of a nasal consonant" (Cruttenden, 2014, p.309) occurs again within words and at word boundaries. Nasalisation of the preceding vowel can be seen in examples such as [n] in *hand*, and also vowel that is between nasal consonants such as in *woman*. Nasalization of the short vowels can be seen in *many*, *funny*, *finger*, and [ɪ] in *strongly*, and also of vowels that follow [m, n], e.g., *meat*. Nasalization at word boundaries sometimes occurs if the boundary of the following word is a nasal consonant, but this type of nasalization can also occur if there is no nasal consonant following.

Cruttenden (2014) also mentions phonemic variations in his work, some of which I already mentioned above as examples of coalescence. There is also the possibility of different phoneme selection of the same word, which depends on the speaker. This can be seen in words such as strength [streŋθ], [strenkθ], and [strenθ].

When comparing the three works of Roach (2009), Cruttenden (2014) and Alameen and Levis (2015), Roach (2009) and Cruttenden (2014) have a similar standing when it comes to the division of connected speech processes. On the other hand, Alameen and Levis (2015) add subparts that are not present in the other two works mentioned or name some of them differently. Their chart involves six categories with multiple subsections, where only one of the main categories is present in Roach (2009). I will not be naming and explaining every single category, but I want to bring attention to at least some of the differences in this research. A category such as "multiple" is absent in Roach (2009). Alameen and Levis (2015) talk about lexical combinations such as *gonna*, *wanna*, and contractions *n't*, *'s*, which are again not

present in Roach. Another interesting point to note is that instead of adding the example of *a bad boy* into the category of modification into the subpart of assimilation, Alameen and Levis (2015) sort it into the category of reduction. When it comes to the category of linking, Alameen and Levis (2015) use the subcategory consonant-consonant with the example *five views*, while Roach (2009) does not recognize this aspect in his research.

### 2.2.2 Elision

Another part of connected speech processes is elision, which often occurs in fast speech. “The nature of elision may be stated quite simply: under certain circumstances sounds disappear. One might express this in more technical language by saying that in certain circumstances a phoneme may be realised as zero, or have zero realisation or be deleted” (Roach, 2009, p. 113).

The first type of elision I am going to mention, is the loss of weak vowels after the consonants p, t, k. When elision of the weak vowel occurs, aspiration of the initial plosive appears, as in *potato* [p<sup>h</sup>'teɪtəʊ], *today* [t<sup>h</sup>'deɪ].

The second type of elision that occurs is when a weak vowel is lost when combined with [n], [l], or [r] and results in a syllabic consonant in words such as *police* [pɪːs] or *correct* [kɹɛkt].

The third type mentioned by Roach (2009) is the avoidance of complex consonant structures. This happens when “word-final clusters of voiceless plosive or affricate + /t/ or voiced plosive or affricate + /d/ (e.g. /-pt, -kt, -tʃt, -bd, -gd, -dʒd/) may lose the final alveolar stop when the following word has an initial consonant” (Cruttenden, 2014, p. 314). Here, the plosive that is in the middle disappears; this can be seen in examples such as *the old man told my brother* [əʊl mæn], [təʊl maɪ 'brʌðə]. The same thing happens when we have a voiceless constituent + [t] or a voiced constituent + [d] followed by a word beginning with a consonant, e.g., *soft rocks, bold friend* [sɒf rɒks, bəʊl frɛnd]. When it comes to the elision of final [t] or [d], it is not so common with initial [h], but there are a few instances, such as in *bald head* or *roundhouse* [bɔːl

həd, raon haos]. Another form of elision also affects final [t] and [d] followed by an initial [j], retain coalescent form as [tʃ] and [dʒ], e.g., *last yell, herded you* [lɑ:stʃ jɛl, 'hɜ:dɪdʒ ju:]. When discussing contractions of grammatical words, it is difficult to decide whether they are affected by elision. Cruttenden (2014) involves them in his research without questions, but Roach (2009) takes a more hesitant stand. Examples such as *not* [nɒt], spelled n't or *is* [ɪz], spelled 's are the most common ones when it comes to their pronunciation after elision takes place. I can also add *have* [hæv], which is often spelled 've. When it comes to negative contractions, ending with [t] is elided when standing before a consonant, *I couldn't stand it* [aɪ 'kʊdn stænd ɪt]. The last type of elision that is talked about in Roach's research is the elision of [v] in *of*, with examples such as *some of them, glass of beer* [sʌm ə ðəm, glɑ:s ə bɪə]. Cruttenden (2014) also mentions a type of elision concerning diphthongs. "When one syllable ends with a closing diphthong (i.e. one whose second element is closer than its first, in GB /eɪ, aɪ, aʊ, əʊ, oɪ/ and the next syllable begins with a vowel, the second element of the diphthong may be elided" (Cruttenden, 2014, p. 314). We can see this on example such as *try alone* [tra ə'ləʊn]. The last part to mention is the elision of initial [ə], mainly when it is followed by a constituent while being preceded by a consonant, e.g., *not another sound* [nɒt n 'ʌðə saʊnd].

### 2.2.3 Linking

The last connected speech process I am going to talk about is linking or liaison. As the main focus of this thesis, this section is going to be the most detailed. I have talked about spontaneous speech and other connected speech processes that influence speech with linking as the focal point of this work. Hieke (1984) talks about the processes that alter speech and classifies linking under the umbrella term of absorption along with levelling and loss. Linking can be consider a subtle form of absorption, where the change in language is not as noticeable as for example loss. These absorption processes create a more fluent sounding language. Hieke (1984) states that, "the derivation of dynamic speech from citation form sets off a range of



absorption processes which are governed, among other things, by the mode of speech involved” (Hieke, 1984, p. 346). He then describes language as "casual and deliberate", which ties into what I have talked about in the part of spontaneous speech. Comparing fluent spoken speech and read, laboratory speech, is where the discussion of connected speech is the most fruitful (Hieke, 1984, p. 346). But what is the reason we use linking so often? Why do we use it? Skarnitzl et al. (2022) state that "the use of linking or glottalization contributes to the characteristic sound pattern of a language, and the use of one in place of the other may affect a speaker’s comprehensibility and fluency in certain contexts” (Skarnitzl et. al, 2022, p. 941). Speakers want to avoid leaving sounds disjoined from each other, filling in undesired gaps in speech. Another aspect why we link is that initial vowels are preceded by a glottal onset and , therefore, are also disconnected. Glottal onset, phenomenon connected to initial vowels, is a sort of interruption in speech, where speaker’s glottis closes and therefore speech is disconnected. Solution for this problem lies in the idea of resyllabification, “where the stress has to be on the initial syllable of the second word for C-V linking to occur” (Alameen, 2007, p. 10). Example of this idea can be seen on the phrase *nicked it*, where /t/ belonging to *nicked* moves to *it*, [nik.tit]. Resyllabification does not always have to take place, mainly when we want to deliberately pause or start speaking again after a silence. In this case we use glottalization.

Moving on to the description of linking, there exist three main types. What this study will focus on, is the type of linking where a consonant links to a following word that begins with a vowel and how vowel endings link to vowel beginnings. The third type is when a consonant ending links itself to a syllable with a consonant beginning. Although it is important to note that this type of linking exists, it will not be discussed in the practical part of this thesis.

Consonant to vowel linking, as mentioned above, consists of a word-final consonant syllable link to a word-initial vowel. This is where resyllabification takes place. In order to

avoid glottal onset before the initial vowel, what happens is that we “move” the final consonant to the initial vowel. Coming back to the example I used previously where /t/ in *nicked it* moves to *it*, [nɪk.tɪt].

Another aspect of linking that belongs in this category is linking /r/. This occurs in non-rhotic accents, in this work it will relate to British English. As Roach (2009) states, “the most familiar case is the use of linking r; the phoneme r does not occur in syllable-final position in the BBC accent, but when the spelling of a word suggests a final r, and a word beginning with a vowel follows, the usual pronunciation is to pronounce with r” (Roach, 2009, p. 115). Linking /r/ is either combined with sounds /ɑ:, ɔ:, ɜ:/ or those that include a final /ə/, such as /ə, ɪə, ʊə/. Examples of linking /r/ can be *far out* or *upholster it*. There are certain conditions when the possibility of linking /r/ is more probable to occur. If a following word begins with a vowel, linking /r/ is present, although not obligatory, e.g. *hear it*. (Cruttenden, 2014)

Moving on to the description of vowel to vowel linking, this type of linking occurs when a word, ending with a vowel, is linked to the following one that begins with a vowel. They are linked by the so called transient /j/ or transient /w/. The rule, whether to use transient /j/ or /w/ is as follows. If a word ends with a high front vowel, /i:, ɪ, eɪ, aɪ, ɔɪ/, then we insert a transient /j/, e.g. *my ears* [maɪ ɪəz]. If a word ends with a high back vowel, /ʊ, u:, əʊ, aʊ/, then the following word is linked with a transient /w/, e.g. *you are* [ju ˈɑ:]. (Alameen, 2007; Cruttenden, 2014). What is interesting to note is that Hieke (1984) does not use the terms transient /j/ or /w/, instead he uses something called “glide-attraction”. He states that, “word-final vowels having offglides (i.e., high and mid vowels) may become linked to the following syllable by their offglide, if that syllable begins with a vowel” (Hieke, 1984, p. 351).

Coming back to vowel to vowel linking, I would like to mention a phenomenon called intrusive /r/. Linking /r/ in British English occurs in vowel-to-vowel linking and is, in some cases, called intrusive /r/. Not so long ago intrusive /r/ was perceived as an undesirable and unaccepted

aspect of speech but current research views it differently. Although the following part is not relevant to current studies anymore, it is interesting to note the changes that this phenomenon underwent in the past few years. Acceptable use of linking /r/ occurs when in the word being linked exists /r/ in the spelling. When there is none, we call that an intrusive /r/, which is, in most cases, an undesired feature. When justified, linking /r/ needs to be "historical", meaning it has some basis in the previously used spellings. Intrusive /r/ occurs often, mainly when the word ends with a /ə/ sound, e.g. *India and Philippines* [ɪndiə r ən fɪlɪpiːnz]. In instances as these, intrusive /r/ is often unconscious, but if the linking follows /ɑː, ɔː/, it becomes easier to notice the usage of intrusive /r/, e.g. *now draw it* [naʊ drɔːr ɪt] (Cruttenden, 2014). The conditions where the possibility of occurrence of intrusive /r/ in vowel-to-vowel linking is higher are as follows. If we have an /ɑː, ɔː/ sound, the inserted /r/ is intrusive and not acceptable, e.g., *now draw it* [naʊ drɔːr ɪt]. Another possibility for intrusive /r/ is before a suffix, which is also highly unacceptable, e.g., *drawing* [drɔːrɪŋ] (Cruttenden, 2014).

The last type I will talk about that is sometimes considered as the third type of linking is consonant-to-consonant linking. As mentioned above, this work will not analyze this type of linking, but I believe it is also important to introduce. There are two things that can happen when two consonants meet at a word boundary. If a word ending in a consonant syllable links to an initial consonant, which is the same, then the consonant sound is prolonged, e.g., *but to* [bətːo]. If, on the other hand, the consonants meeting at a word boundary are different, then the final consonant is moved to the following word, e.g., *what sheep* [wətʃiːp] (Alameen, 2007). What is interesting to note is that Hieke (1984) labels consonant-to-consonant linking as "release-attraction". He states that, "the trend to avoid hiatus extends to consonants if followed by continuants, with the aspiration that is part of the release portion of that stop actually remaining unreleased until production of the following continuant" (Hieke, 1984, p. 352). He uses the

examples as "what she [wa.t<sup>o</sup> ʃiy]" and "about the [a.bau.t<sup>o</sup>ð]" where the final consonants are marked with a non-release.

Another form of release attraction, "nasal-release", involves nasals, specifically sequences /tn/ and /dn/. In a word as goodness, [dn] sounds is moved to the second syllable of the word, in the first it is silenced, [gʊd<sup>~</sup> nəs] (Hiecke, 1984).

The last thing I would like to mention is the phenomenon of juncture, which is closely connected to linking. Roach (2009) ties it in with linking and intrusive /r/, while Cruttenden (2014) refers to it in means of word or morpheme boundaries. In language there exists something called minimal pairs. Those are words that are different in only one sound. When combined with one more word and contrasted with another minimal pair, we can see how word boundaries work and how they affect the sound production. Minimal pairs help us understand the importance of juncture in language (Roach, 2009; Cruttenden, 2014). We can see it on an example, *great ape* and *grey tape*. Both are transcribed /greɪtɛɪp/, but they have a different word boundary, as can be seen here [greɪt ɛɪp] and [greɪ tɛɪp]. This is where juncture helps us understand what marks the difference in a minimal pair. In the example *grey tape* [greɪ tɛɪp], the difference is that the /t/ in *tape* is fully aspirated, and the /eɪ/ in *grey* is pronounced with a full length. On the other hand, in *great ape* [greɪt ɛɪp], the pronunciation of /eɪ/ in *ape* is a lot shorter, while /t/ in *great* is only slightly aspirated (Cruttenden, 2014; Roach, 2009).

### **3. Materials and methodology**

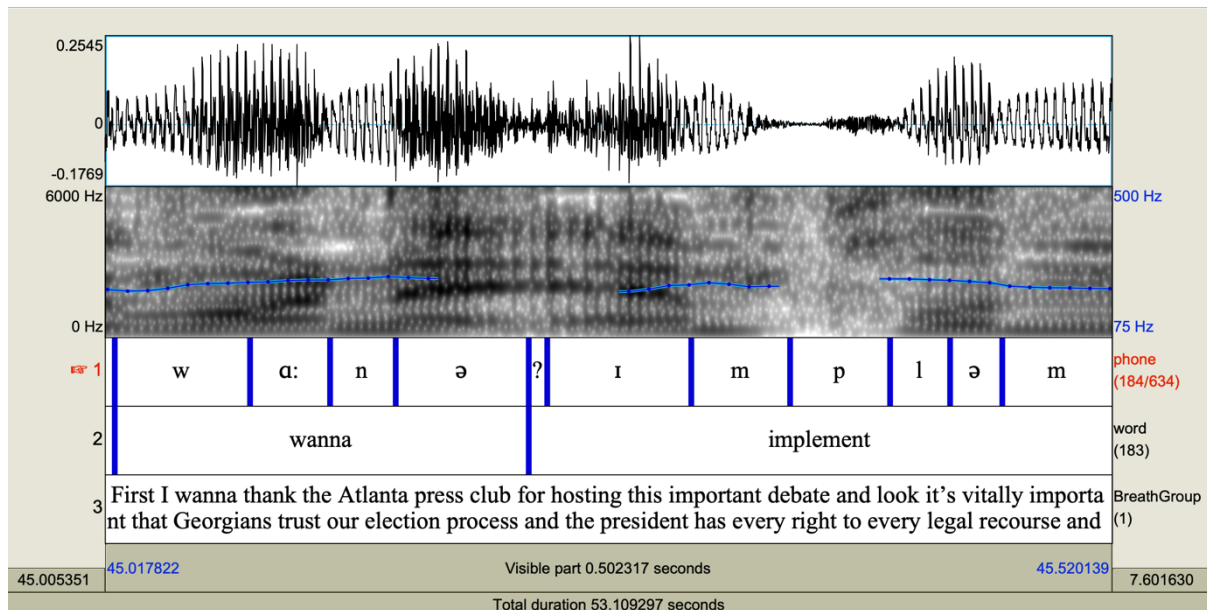
As mentioned in the previous parts, this thesis concerns linking in spontaneous speech. The practical part of this thesis has been done in a computer program called Praat, which is used for phonetic research.

The materials used for this research are 16 tape recordings of British and American speakers, which were recorded and uploaded to the internet, where they are freely accessible. These recordings are 1-to-2-minute snippets of political debates. Of the 16 recordings, 8 were in American English, and 8 were in British English. For this research to be accurate, we chose 8 female and 8 male speakers, 4 of whom spoke in American English and 4 of them spoke in British English.

Using the textgrid and wav. files of the recordings uploaded into Praat (Boersma & Weenik, 2024), I started adjusting the word boundaries and focused on the aspect of linking. Since this research is interested in linking, I focused mainly on the vowel sounds, whether it came to adjusting the word boundaries or the linking itself. Going through the process, I analyzed every word looking at the spectrogram which showed where the sound boundaries could be. Since I hadn't done this type of research before, I had to be extremely careful, closely listening to every word so the boundary would be in its correct place. After analyzing a few recordings, it became easier and easier to decide where the sound boundary would be, since the combination of sounds is not endless and started to repeat. There were some problematic areas that I needed to consult but I will say more about that in a later section.

When it came to the linking itself, I created a document where I pasted the text that I was analyzing and marked every single instance of linking to check if I missed any in the Praat. As mentioned before, this thesis is concerned with V-V linking and C-V linking only, leaving out C-C linking, which is sometimes also treated as linking. Some instances were harder than others, while others were clearly visible, not needing to listen to that part of the recording. I

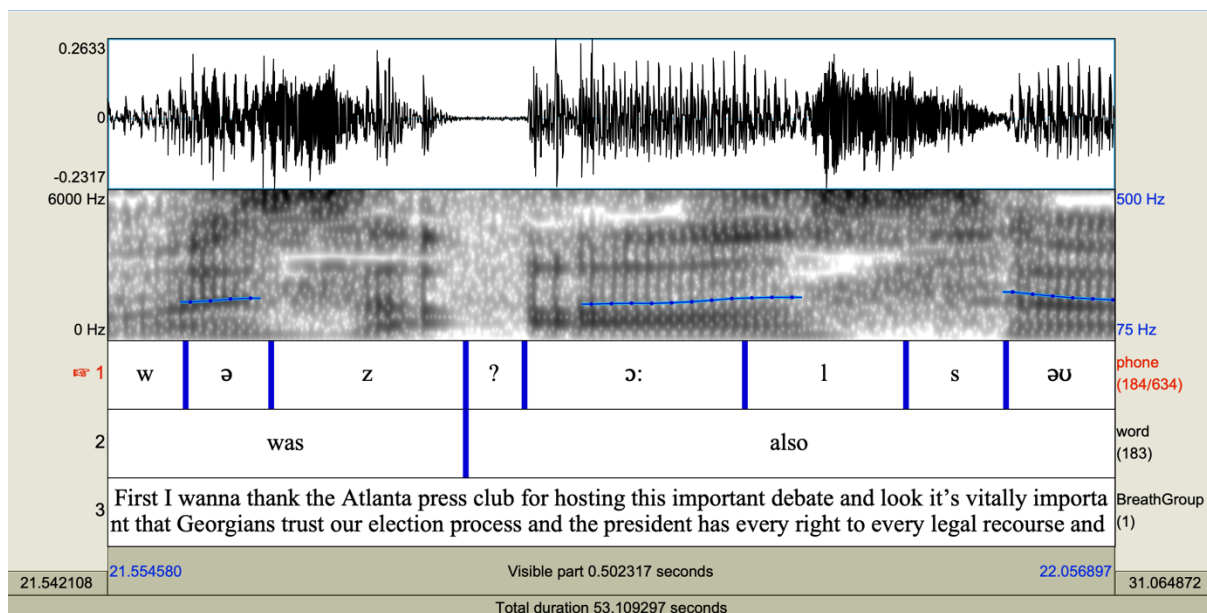
will use a few examples that are going to show the instances that were a bit harder to decide upon and those that were, on the other hand, very easy.



**Figure 1.** Occurrence of glottalization in the American data.

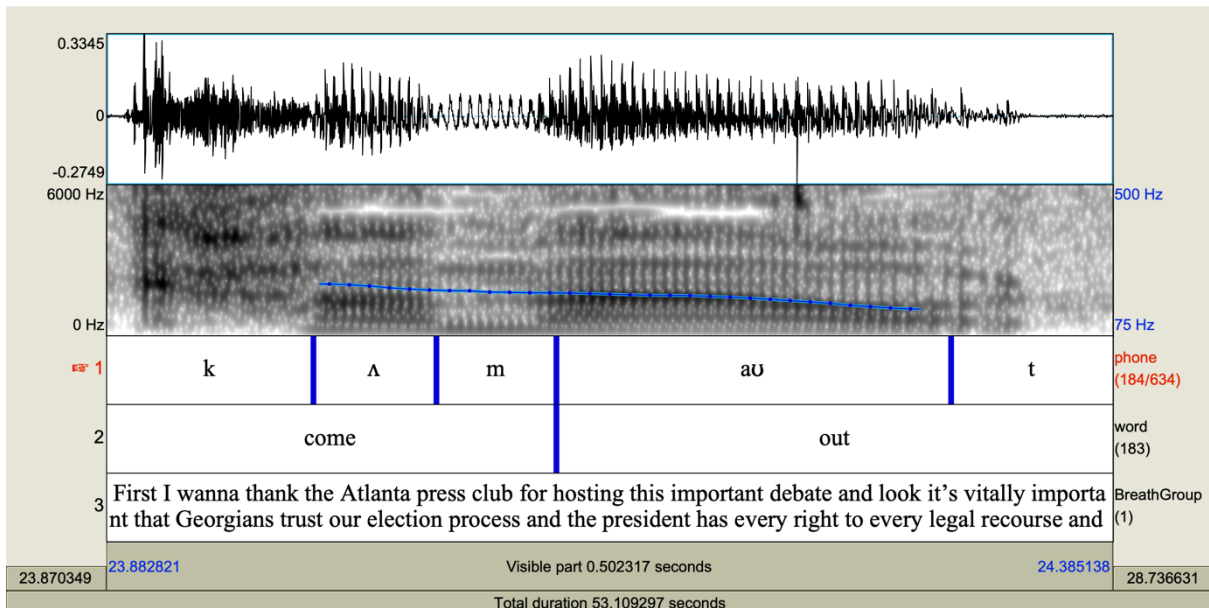
The first example I am going to use is an example with glottalization. In this example and one that will follow is important to mention that the boundaries of glottalization are not one hundred percent exact and the important thing to notice is that glottalization is present. As can be seen in **Figure 1** we are dealing with the question whether the words *wanna* and *implement* are linked or glottalized. In cases like this it is not so clear, and we cannot decide by just looking at the spectrogram whether one or the other is taking place. In the following example, **Figure 2**, we will be able to see that glottalization can be in some cases obvious. In **Figure 1** we cannot be one hundred percent sure since the vertical and horizontal lines that connect the words *wanna* and *implement* are present, showing a continued sound. Here, we have to do a close listening, ideally form a word that would capture the sound these two words would make if they were linked. I will use an example to make this understandable. If I have words *police* and *are*, and they are linked, the sound that I will “create” from these two to check if they are linked in a recoding will be [pɒli:sɑ:] since it represents how *police* and *are* would sound merged together. Repeating this word a few times out loud helps immensely when

making this decision. After this, we play the part of the recording we are analyzing and try to decipher whether our made example sound is produced or if we hear glottalization taking place.



**Figure 2.** Occurrence of glottalization in the American data.

In this second example we can see another instance of glottalization. **Figure 2** has a straightforward answer, we do not have to think long before deciding whether *was* and *also* are going to be linked or glottalized. In comparison to **Figure 1**, here the vertical and horizontal lines are not present which we can assume means a small pause, therefore making *also* glottalized. In cases like this we do not need to create an example sound of the two words being linked nor do we need to closely listen to this part of the recording. The spectrogram shows us the answer straight away.



**Figure 3.** Occurrence of linking in the American data.

The last example I will use is an instance of C-V linking, as seen in **Figure 3**. The process of deciding whether linking or glottalization is occurring is the same as in the first example. First, we create an example sound of words *come* and *out* linked, then we proceed with playing the recording until we decide upon whether these two words are linked or glottalized. Some cases are easier than others; in this case, we can see the vertical and horizontal lines are uninterrupted, so we can assume that this will be the case of linking.

When analyzing spontaneous speech, there are instances where the speaker makes certain mistakes, making the analysis somewhat more complex but also more dynamic than, for example, the analysis of laboratory speech. During my analysis, there were some problems when it came to the content of those recordings, such as hesitations, disfluencies. Some of the speakers were unsure during their speech resulting in more hesitations but there were also instances of disfluencies that disrupted the flow of speech making it more difficult to analyze. It is important to mention that all disfluencies, hesitations and pauses were not included in the results. Another challenging aspect of this analysis was the fast speech of some speakers. If the speaker spoke fast the adjustment of sound and word boundaries became very difficult. In these cases, there were more instances of left-out sounds, making the word sound very different than



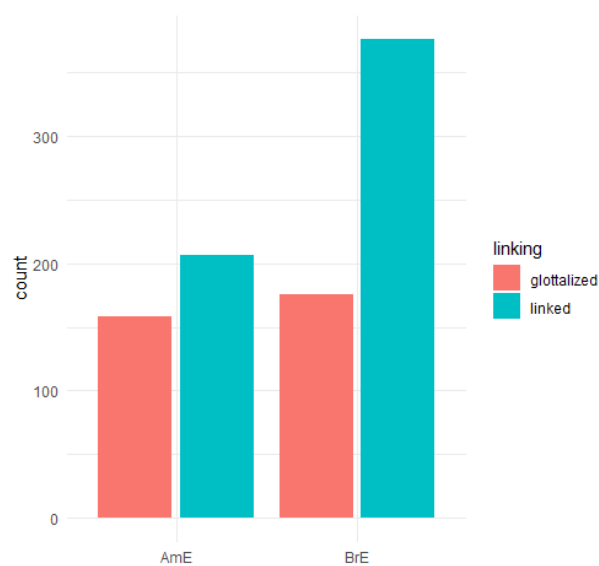
its standard form. In these cases, I had to make adjustments, delete certain phones, and decide whether the word was helpful in the analysis or not. As mentioned above, some speakers were either nervous or unsure of the content of their speech, which affected the amount of linking or glottalization used in the recording.

After the analysis of all 16 recording tapes was finished, the results were put through a Praat script to extract the data. Using Excel to sort out the results, e.g., marking all the lexical words as lexical and all the grammatical words as grammatical and deleting instances of disfluencies and hesitations that could influence the final results, and then R Studio was used to process and visualize the data using the ggplot2 package (Wickham, 2009).

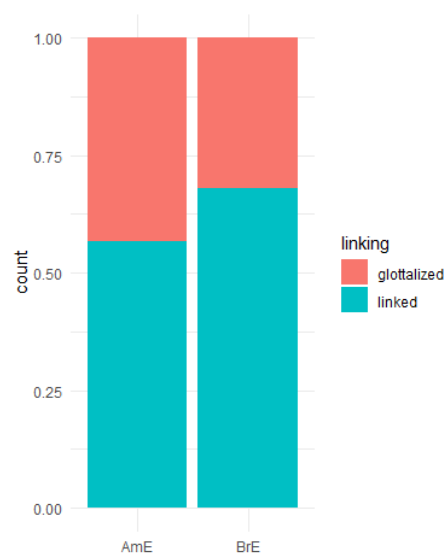
#### 4. Results and discussion

This section will focus on the results and discuss them. The discussion will focus on the general and individual tendencies of speakers. In most examples, I will use two graphs showing the same result so we can better understand the results.

The first two graphs will focus on the variety of language with a focus on the aspect of linking.



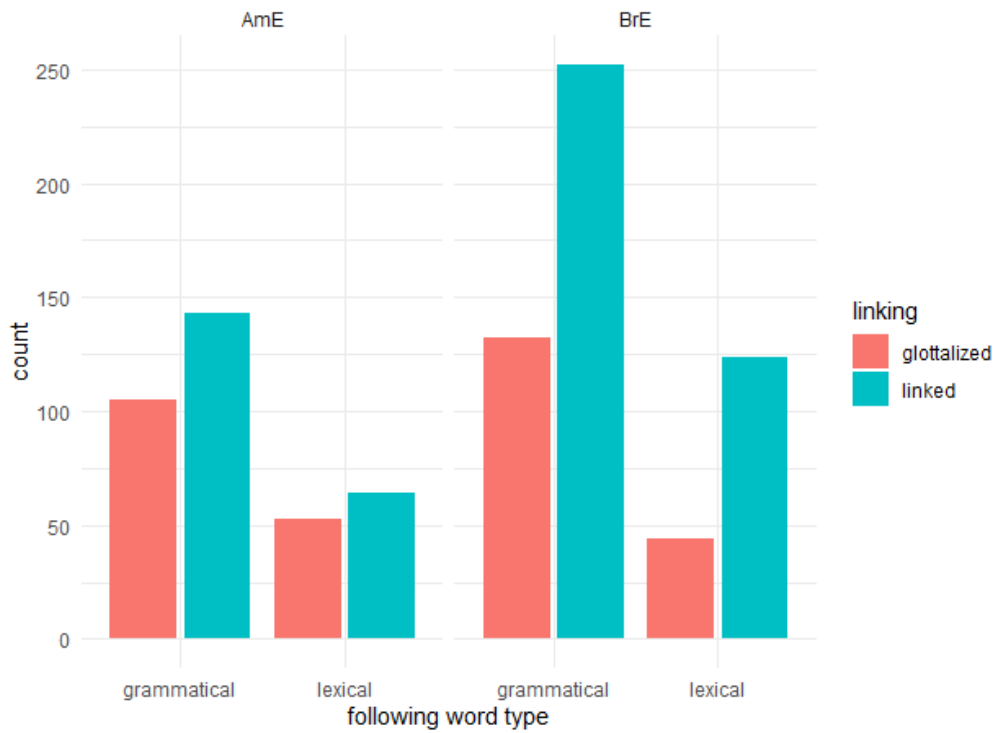
**Figure 4.** Bar chart showing the overall number of linking vs glottalization for speakers of American and British English.



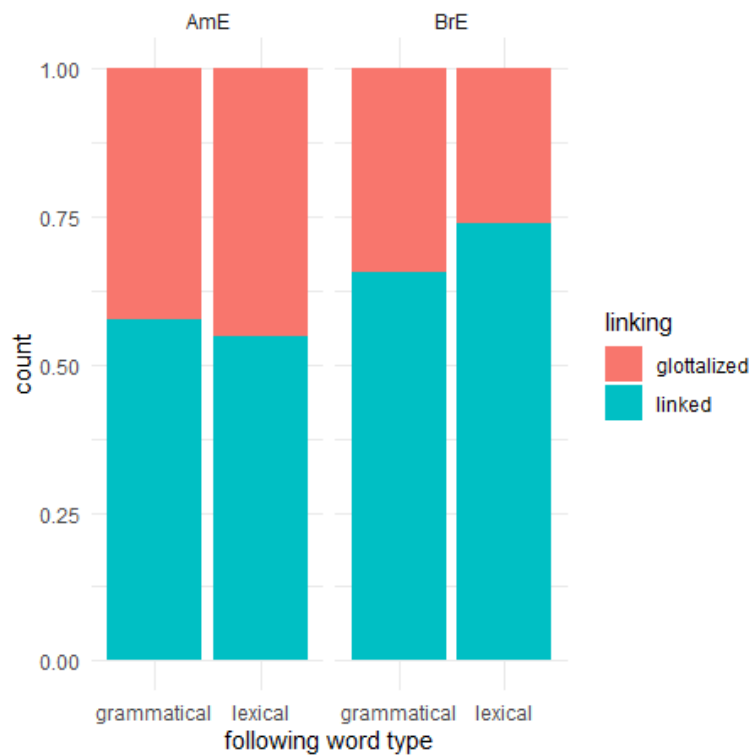
**Figure 5.** Percentage bar chart demonstrating the distribution of linking and glottalization for speakers American and British English.

In **Figure 4**, we can see the comparison between American and British English with a focus on the number of occurrences. In **Figure 5**, the focus is on a percentage rather than the precise number. These two graphs show that British English speakers are more prone to linking than American speakers, while American speakers have the difference balanced more evenly. Even though the difference in **Figure 5** is not as meaningful, it does show us a certain preference on the side of British speakers. On the other hand, it is important to look at **Figure 4**, which shows us that the number of linking occurrences was higher in British English than in American, with the number exceeding 350 occurrences. This is going to occur throughout almost every graph, the reason being that British speakers had a higher word count and faster speech than American speakers. When looking at all of the results, this seems to be the most promising answer to why the count in British speakers is always significantly higher than in the count of American speakers. All recordings had a similar length, whether it was an American or British speaker, but the speech rate was significantly different when it came to the language variety.

As mentioned above, one reason why British speakers had a higher count of linking and glottalization is speech rate. Some of the speakers of American English were talking slowly in their speech or sometimes with emphasis on certain words. Some of them had a combination of fast speech and some of them were taking more pauses during speech. The British speakers on the other hand spoke with a faster speech rate, resulting in a bigger word count and a higher possibility of linking.



**Figure 6.** Bar chart showing the overall number of linking vs glottalization for speakers of American and British English per word type.



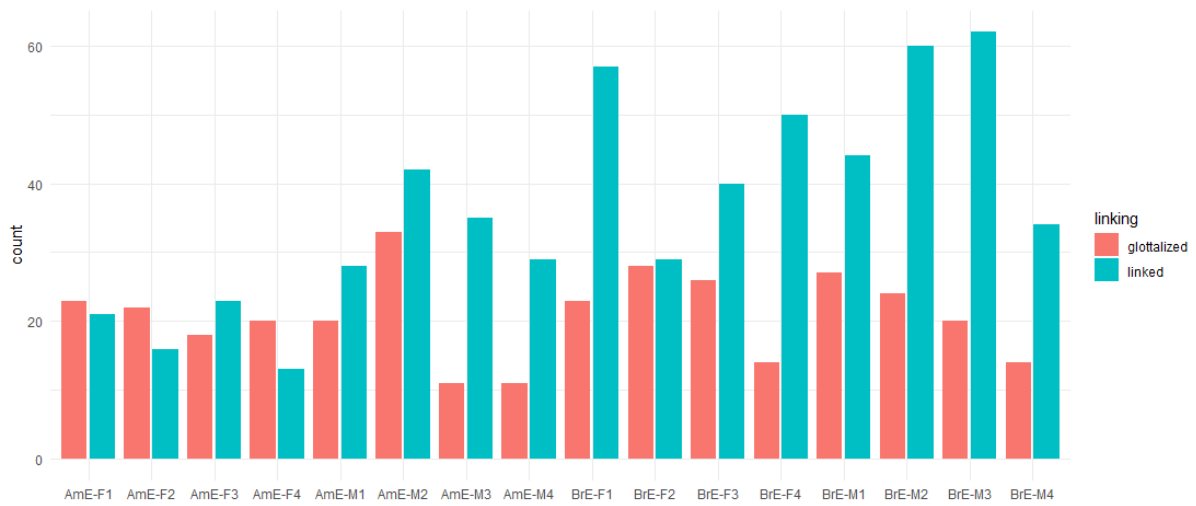
**Figure 7.** Percentage bar chart demonstrating the distribution of linking and glottalization per word type.

In **Figure 6** and **Figure 7**, we can see the results with information regarding variety and word type. As we saw in **Figure 4** and **Figure 5**, the number of occurrences is again higher in British English than in American English. The reasoning behind this would be the same as in previous discussion that faster speech rate affected the results of this group of speakers.

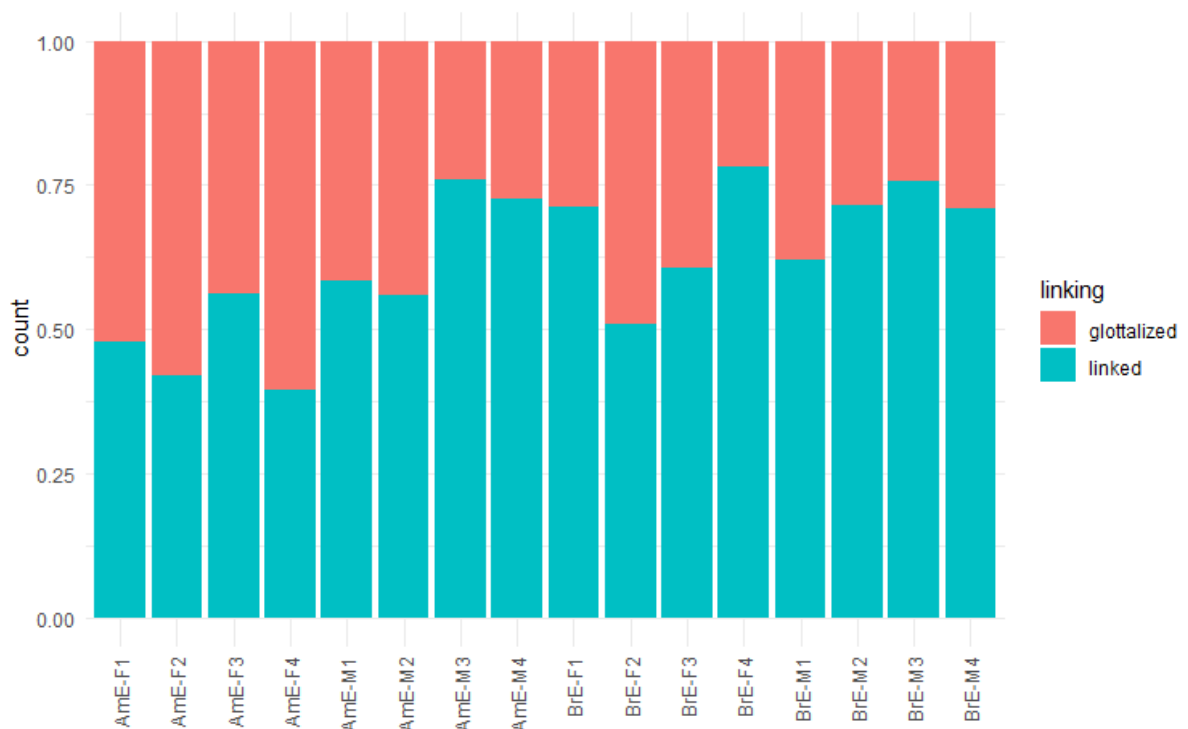
What is interesting in these results is the fact that when comparing it to the existing phonetic research, it seems that these results have shown us the opposite of what we are told about linking in speech. Generally, grammatical words are supposed to have a higher linking count than lexical words, meaning we link grammatical words significantly more than lexical ones. Speakers are then more likely to link words *and up* than *and open*. What is important to mentioned is that the information we have from researchers relates to laboratory speech more. Spontaneous speech has not been extensively researched, mainly in the area of connected speech processes, so the knowledge we have regarding speech is not useful when applying it to spontaneous speech. When we look at the two graphs, we can see that speakers of American English have a higher linking count of grammatical words than lexical. Since this research is not extensive and we do not have the amount of data that we would need to make a convincing argument that American speakers adhere to the "rules" produced by the research of laboratory speech, I am going to leave this part open. The amount of linking of grammatical and lexical words in American English may be therefore balanced, higher or lower. Only further, deeper, research could tell us more about this result.

On the other hand, British speakers link lexical words more frequently than grammatical ones. This disproves the rules produced by the research of laboratory speech but again, the same question arises with the validity of these results since this research is not extensive. Interesting thing to add is that American speakers, in most cases, did not link the grammatical word *and*, instead, they used it as an opportunity to take a pause. We decided not

to include these instances in the results, since it is clear that glottalization would be present in every occurrence.



**Figure 8.** Bar chart showing the overall number of linking vs glottalization for individual speakers.



**Figure 9.** Percentage bar chart demonstrating the distribution of linking and glottalization for individual speakers.

**Figure 8** and **Figure 9** show us the results of tendencies of individual speakers. Here I am going to compare two American speakers and two British speakers with the biggest differences between them, and then do overall comparison of two extreme cases between American and British speakers.

Firstly, let us look at the values of AmE-F4 and AmE-M3. In **Figure 9** speaker AmE-F4 has the count of linking compared to glottalization only around 38 percent. As I listened to the specific speaker again, I noticed the reason that caused this result. It is the same reason that tends to influence all of the results we see in this thesis. The female speaker has a very slow speech rate that caused linking to diminish. On the other hand, speaker AmE-M3 had a fast speech rate, the flow of the speech was a lot more dynamic causing linking to be more frequent. Moving on to British speakers, where the biggest difference can be seen between BrE-F2 and BrE-M3.

In **Figure 8**, the female speaker BrE-F2 has the most balanced linking and glottalization count. When comparing it to the male speaker, BrE-M3, he has the most extreme result of all in absolute numbers. This we can also solve with the idea of speech rate. He could have been an extremely fast speaker, or the recording could have been longer, with a higher word count because of fast speech but also because of the length of the recording. Although the time of all speaker group recordings should be similar, when it comes to the individual recordings they may differ from 52 seconds to 92 seconds per recording. When comparing this through variety, two extreme cases, as mentioned above, are AmE-F4 and BrE-M3. I am not going to explain this in detail again, but the answer to why the difference between these two speakers is so great is speech rate.



**Figure 10.** Percentage bar chart demonstrating the distribution of linking and glottalization per word type for individual speakers.

**Figure 10** shows both variety and word type linking preferences of individual speakers.

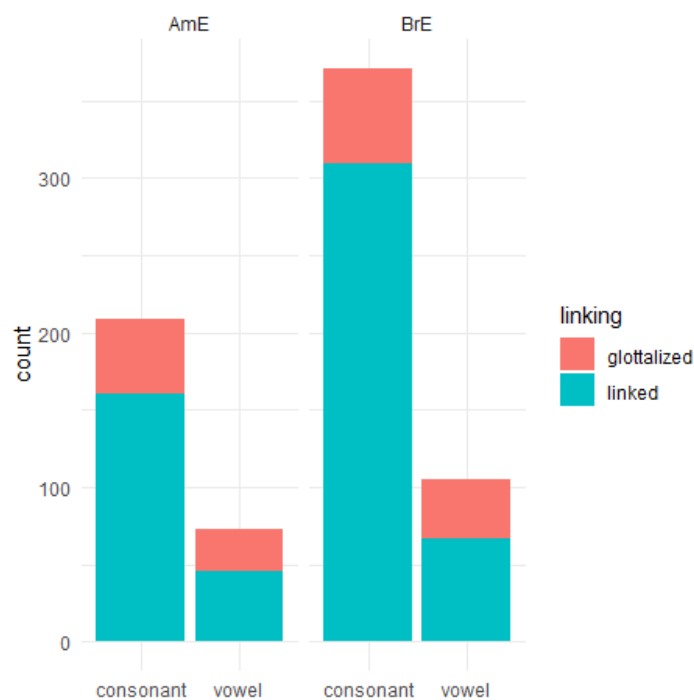
I am going to pick extreme cases again to show the differences between individual speakers. Let us first look at the grammatical category. In speaker AmE-F2 the percentage of linking is only around 36 and is the lowest amount that we have in this category. Similar to the cases before, this female speaker had speech rate on the slower side, using a lot of emphasis.

On the other hand, cases like AmE-M3, AmE-M4, or BrE-M3 show a higher count of linking grammatical words, each with 75 percent. This can be again argued with the problem of speech rate. Moving to lexical category, AmE-F1 shows the lowest percentage of lexical words linked. It is interesting to note that the speaker linked grammatical words with a count of over 50 percent.

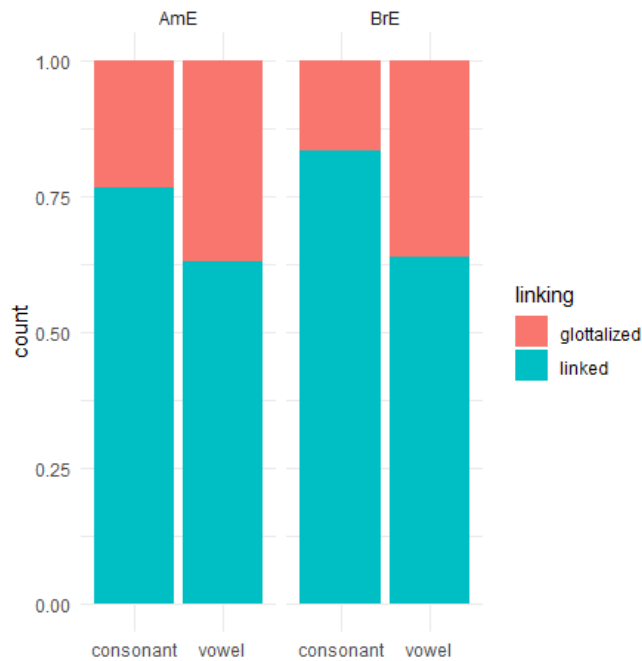
If we come back to the **Figure 8** and **Figure 9**, there the results seemed to be balanced when it came to glottalization and linking. Here, we have it divided into word type categories, and the preferences of individual speakers may be different than when grouping them together as a variety. Another interesting point is that in this graph we cannot see if the number of occurrences between grammatical and lexical words is the same or if one category has a more



prevalent count than the other. Secondly, the lexical word type category also shows some interesting results. When we compare AmE-F1 and BrE-F4 the difference between these two is highly significant. The American female speaker linked lexical words only about 25 percent from all of the occurrences. The British female speaker linked lexical words over 90 percent from all the occurrences and her amount of linking of grammatical words was also quite high. This significant difference can be once again explained by the use of speech rate. As I mentioned before, American speakers spoke a lot slower, although not all of them but the majority did which affects the role and amount of linking that occurs in speech. When we transform this point to the opposite side, we get the answer why does the British speaker link so frequently.



**Figure 11.** Bar chart showing the overall number of linking vs glottalization for speakers of American and British English specifically focused on consonants and vowels.



**Figure 12.** Percentage bar chart showing the overall number of linking vs glottalization for speakers of American and British English specifically focused on consonants and vowels.

**Figure 11** and **Figure 12**, show the count of consonant and vowel linking with a focus on language variety. As mentioned above in chapter 3, these results and all of those that were already discussed are without existing pauses in the recordings. When there is a pause, vowel is automatically glottalized, so including pauses in these figures would heavily influence the outcome.

In **Figure 11** on the left side of the graph we can see that in British English the count of occurrences of consonants in the British recordings that could have been either linked or glottalized, is over 350, meaning that the number of instances in all of the recordings, where consonants were either linked or glottalized was over 350, while in American English the count is only slightly over 200. Again, I assign this difference to the speech rate. Since British speakers had a faster speech rate, the count of occurrences of consonants but also vowels that could have been linked or glottalized, is going to be naturally higher, meaning that since the speech rate is higher the word count is also going to be higher leading to a higher the number of vowels and consonants present in all of the British English recordings.

If we look at the count of vowels, this number seems to be a bit more balanced, with British speaker count of slightly over 100 and American speaker count around 75 occurrences. When looking at **Figure 12** the difference does not seem as drastic as in **Figure 11**. The linking of vowels seems to be balanced while the linking of consonants seems to be more frequent in British English. Another interesting thing to note is that both American and British speakers tend to prefer consonant to vowel linking more than vowel to vowel linking.

## **5. General discussion and conclusion**

The research that was conducted gives us an idea of how linking in spontaneous speech really works. This research provides insight into whether the rules for linking, provided by the research on laboratory speech, can also be applied to spontaneous speech. It also provides a broader understanding of the tendencies for the use of linking in British and American speakers.

We know from phonetic research that vowel-to-vowel linking should be more frequent than consonant-to-vowel linking. The results from this research show otherwise. Regardless of dialect, the phonetic tendency of both American and British groups preferred consonant-to-vowel linking. The main aspect that influenced linking in the analyzed recordings was speech rate, which can be applied to all 16 American and British English speakers. The results showed that British speakers have a faster speech rate, resulting in more cases of linking present in the recordings. In comparison, American speakers had a tendency for a slower speech rate, resulting in fewer instances of linking.

Regarding speech rate, I want to come back to the studies by Kohler (1996) and Trouvain et al. (2001), where they focused on an articulation rate in spontaneous speech and found out that spontaneous speech has higher speech rate and greater variability than read speech. This result directly correlates with what was observed in the results of this thesis.

Another research mentioned in the theoretical part relevant to the analysis of recordings were studies conducted by Johnson (2004) and Dilts (2013), which targeted the issue of variability in speech. Both of them investigated how common is canonical form of word in speech and the amount was that 32 percent of words did not occur in their canonical form. During my analysis in the practical part of this thesis it became clear that canonical form was not frequently present in some words, e.g. *and*, occurred most commonly in the form of [ən].

It is difficult to make generalizations about linking in spontaneous speech. We can divide speakers' tendencies into individual and general, but it is difficult to strictly decide whether the

common tendencies can be applied to all. For this research to tell us more about the general and individual tendencies for linking in English, we would need to conduct a deeper study with more data to provide us with more information.

As mentioned above, for future research it would be productive to collect more data so the results could tell us more about linking in spontaneous speech. Further research in this area could provide more conclusive results regarding the tendencies of linking. Further research of this area would be beneficial for L2 learners' ability to learn how to sound more native and fluent in their speech.

## 6. References

- Alameen, G. (2007). The use of linking by native and non-native speakers of American English. *Retrospective Theses and Dissertations*. Iowa State University.
- Alameen G. & Levis, J. M. (2015). Connected Speech. In: Reed M. & Levis M. J. (eds.), *The Handbook of English Pronunciation*, 159-174. Oxford: Blackwell.
- Boersma, Paul & Weenink, David (2024). Praat: doing phonetics by computer [Computer program]. Version 6.4.16, retrieved 29 July 2024 from <http://www.praat.org/>
- Cruttenden, A. (2014). *Gimson's Pronunciation of English*. Routledge.
- Hieke, A. E. (1984). Linking as a Marker of Fluent Speech. *Language and Speech*, 27 (4), 343-354.
- Oppenheimer, D. M. (2008). The Secret Life of Fluency. *Trends in Cognitive Sciences*, 12 (6), 237-241.
- Roach, P. (2009). *English Phonetics and Phonology*. Cambridge University Press.
- Skarnitzl et al. (2022). Glottalization and linking in the L2 speech of Czech learners of Spanish, Italian and Portuguese. *Second Language Research*, 38(4), 941-963.
- Tucker, B. V. & Mukai, Y. (2023). *Spontaneous Speech*. Cambridge University Press.
- Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer

## 7. Resumé

Hlavnými cieľmi tejto bakalárskej práce bolo zistiť či pravidlá viazania reči, ktoré vychádzajú z výskumov robených v laboratórnych podmienkach sa dajú aplikovať na spontánnu reč a porovnať spôsob viazania v Americkej a Britskej angličtine. Zamerali sme sa najmä na porovnanie počtu viazania a glotalizácie v nahrávkach, ako aj jednotlivu na slovné druhy, pohlavie alebo na jednotlivých hovoriacich.

Teoretická časť práce sa začína v 2. kapitole. Tu sú vysvetlené rôzne aspekty reči, ktoré ju ovplyvňujú, s dôrazom na spontánnu reč, ako aj príklady štúdií, ktoré sú relevantné k porozumeniu spontánnej reči ako takej. V úvode kapitoly je vysvetlené čo je to plynulá reč a ako interpretujeme informácie. Plynulosť je mentálny proces, ktorý používa každý z nás. Je to naše chápanie určitých vecí, a hlavným cieľom je zistiť, či je tento proces ľahký alebo ťažký. Náročnosť tohto procesu sa dá taktiež manipulovať. Ak chceme proces urobiť ťažším môžeme napríklad zmenšiť text v knihe na ťažko čitateľný. Ak ho chceme uľahčiť, môžeme napríklad rozprávať pomaly a zreteľne, aby nám druhý ľahko rozumel, a tým pádom vedel rýchlo spracovať informáciu, ktorú mu dávame. Táto kapitola sa zaoberá najmä tým čo je to spontánnu reč. Spontánnu reč je ťažko definovateľný termín. Každý z nás môže považovať za spontánnu reč niečo odlišnejšie. Samozrejme sa pohybujeme niekde v rovnakom rozmedzí, no špecifiká bývajú občas rôzne. Spontánnu reč je produkovaná v neformálnom prostredí, je neplánovaná, dynamická a nie je čítaná. Je taktiež dôležité zmieniť, že výskumy spontánnej reči nie sú také rozšírené ako výskumy reči robené vo fonetických laboratóriách, v sterilnom prostredí, kde sa spontaneita reči nevyskytuje, je teda ťažké aplikovať výsledky týchto výskumov na spontánnu reč.

V podkapitole 2.1.1, 2.1.2 a 2.1.3 sú rozobraté aspekty, ktoré ovplyvňujú spontánnu reč. Týmito aspektami je variabilita, predvídateľnosť a prozódia. Variabilita reči hrá veľmi dôležitú úlohu v spontánnej reči. Spontánnu reč je sama o sebe plná variability, hlavne čo

sa týka slov. Slová majú svoju slovníkovú podobu, no zároveň aj mnoho iných podôb, ktoré využívame v spontánnej reči. Variabilitu ovplyvňujú procesy ako napríklad skracovanie, zmazanie alebo neukončené vyslovenie. Slovníková podoba slov sa môže nachádzať najmä vo veľmi pomalej čítanej reči, ktorá je presným opakom spontánnej reči. Variabilita sa spája s témou kontextu v reči. Kontext môžu byť sociálne faktory, prostredie alebo okolitý zvuk. Ak sa bavíme o variabilite musíme zároveň zmieniť kontextové podmienky, ktoré ju ovplyvňujú. Prvou kontextovou podmienkou je predvídateľnosť. Táto kontextová podmienka nám pomáha predvídať reč predvídaním koľko redukovania sa bude vyskytovať v konverzácií ak je dĺžka slov kratšia. V tejto časti sú opísané rôzne výskumy, ktoré sa zaoberajú tematikou predvídateľnosti. Druhou kontextovou podmienkou je prozódia, ktorou taktiež dokážeme predvídať spontánnu reč. Prozódia sa zaoberá časom trvania, vďaka tomuto dokážeme predvídať či reč, ktorú počujeme, má črty spontánnosti. Ak je teda tempo reči rýchle vieme predvídať, že reč bude skôr spontánná. Poslednou podkapitolou tejto časti je podkapitola 2.1.2, ktorá sa zaoberá zmenou štýlov reči. Štýl reči môže byť napríklad nenútená alebo čítaná reč. Ak je niekto účastníkom nenútenej konverzácie, v ktorej sa cíti komfortne, redukcia slov je veľmi vysoká. Naopak pri konverzácií formálnej alebo opatrnej budeme vidieť presný opak. Hovoriaci si bude dávať pozor na slova, ktoré vyberá a ich výslovnosť.

Podkapitola 2.2 už priamo hovorí o spojenej reči, na ktorú sa táto práca sústreďí. Spojená reč znamená, že ak plynule rozprávame, nejaké slová, ich začiatky a konce, budú do seba kĺzať. Každý z nás používa spájanie v reči, ak by sme ho nepoužívali, zneli by sme veľmi neprirodzene. Máme niekoľko procesov, ktoré zaradzujeme do spojenej reči, sú to asimilácia, elízia a viazanie. Asimilácia je proces spojenej reči, ktorý ovplyvňuje to ako sa slová, ktoré sa nachádzajú vo svojej blízkosti vyslovujú. Asimilácia ovplyvňuje najmä samohlásky. Asimiláciu rozdeľujeme do dvoch kategórií, regresívnu a progresívnu.



Regresívna asimilácia mení foném predchádzajúceho slova na foném nasledujúceho slova. Progresívna asimilácia je presný opak, foném nasledujúceho slova sa mení na foném predchádzajúceho slova. Asimilácia sa taktiež rozdeľuje do asimilácie miesta, spôsobu a znelosti, a ďalej teda na regresívnu a progresívnu.

Elízia je ďalší proces spojenej reči. Nachádza sa často v rýchlej reči a môžeme ju definovať ako proces počas, ktorého dochádza k strate zvukov. Môže dochádzať k strate slabých samohlások po spoluhláskach p, t, k. Môže dochádzať k strate slabých samohlások ak sú kombinované s [n, l, r]. Posledný typ elízie spôsobuje vyhýbanie sa komplexnej spoluhláskovej štruktúre. To znamená, že pri určitých afrikátoch alebo nezvučných plozívach v spojení so zvučnými plozívami, afrikátmi alebo /t/ môže dôjsť k strate finálnemu alveolárnemu zastaveniu ak nasleduje slovo so začiatočnou spoluhláskou.

Podkapitola viazania je v tejto teoretickej časti skoro najdôležitejšia, keďže opisuje pravidlá viazanie, ktoré sú neskôr skúmané v praktickej časti. Viazanie môžeme chápať ako formu absorpcie zvukov, kde zmenu vo výslovnosti nezaznamenávame až tak veľmi ako napríklad u elízie. Procesy absorpcie zvukov spôsobujú to, aby reč znela kĺzavo, plynule. Poznáme tri typy viazania. Prvý typ je viazanie je viazanie spoluhlásky ku samohláske. Tento typ viazania nazývame aj resilabifikáciou. Aby sme sa vyvarovali glotalizácií samohlásky potrebujeme predchádzajúcu spoluhlásku v slove presunúť k nasledujúcej, začiatočnej samohláske. Pri tomto type viazania používame aj takzvané viazané /r/, ktoré sa používa v britskej angličtine. Ak sa nám slovo končí na spoluhlásku r, automaticky sa prenáša do nadchádzajúceho slova začínajúceho sa na samohlásku. Druhý typ viazanie je viazanie samohlásky so samohláskou. Tu môžeme viazať buď s prechodným /j/ alebo /w/. Ak sa nám slovo končí na vysokú, prednú samohlásku, vkladáme prechodné /j/. Ak sa slovo končí na vysokú, zadnú samohlásku, vkladáme prechodné /w/. Posledný typ viazania, ktorý poznáme, no nie vždy sa uvádza ako viazanie,

je viazanie spoluhlásky so spoluhláskou. Pri tomto type viazanie sa spoluhláska predĺži aby sa spojila s druhou alebo sa posledná spoluhláska prenesie do nasledujúcej spoluhlásky. Tento typ viazania v praktickej časti rozoberať ale nebudeme.

Kapitola číslo 3. hovorí o materiáloch použitých na analýzu a následnú metódu analýzy. Vybrali sme 16 nahrávok anglicky rodených hovoriacich, 8 z nich boli hovoriaci americkej angličtiny, zvyšných 8 bolo hovoriacich anglickej angličtiny. Ďalej sme ich rozdelili do skupín pohlavia kde boli 4 ženy a 4 muži z každej skupiny hovoriacich. Tieto nahrávky boli časti politických debát a každá mala zhruba 1 až 2 minúty. Tieto nahrávky boli spracované v počítačovom programe Praat, kde bolo treba upraviť hranice slov a rozhodnúť či sú segmenty nahrávok kde sa môže nachádzať viazanie zviazané alebo glotalizované. V niektorých prípadoch bolo zreteľné či sa jedná o viazanie, v niektorých to bolo problematickejšie. V tejto kapitole boli taktiež vložené 3 príklady, ktoré ukazujú kde rozhodovanie bolo problematické a kde bolo ľahké. Po dokončenej analýze boli výsledky vložené do skriptu na extrakciu dát. Využitý bol taktiež program Excel na rozradenie výsledkov a R Studio s použitím balíku ggplot2 na vytvorenie grafov.

Nasledujúca kapitola číslo 4. sa zaoberá výsledkami a diskusiou nad nimi. V tejto sekcii bolo využitých 9 grafov na znázornenie výsledkov. Prvý výsledok sa týka rozdielu počtu viazania medzi americkými a britskými hovoriacimi. V grafe bolo možné vidieť že britskí hovoriaci preferovali viazanie viac ako americkí hovoriaci. Aj keď rozdiel medzi preferenciou viazanie a glotalizácie nebol až taký dramatický rozdiel tam bol. Taktiež je podstatné zmieniť to, že vo všetkých grafoch, ktoré sa vyskytujú v tejto kapitole, je počet prípadov kde bolo možné viazať alebo glotalizovať slová oveľa vyšší pri britských hovoriacich ako pri amerických. Dôvod je rýchlejšie tempo reči britských hovoriacich tým pádom aj viac možností kde glotalizovať alebo viazať. Tento fakt bude odpoveďou na väčšinu výsledkov, na ktoré sa prišlo.

Druhý výsledok, ktorý vyplýva z grafov je, že britskí hovoriaci preferujú viazanie lexikálnych slov viac ako gramatických. U amerických hovoriacich je to naopak, preferencia viazania sa nakláňa skôr ku gramatickým slovám. Čo je ale zaujímavé podotknúť je, že podľa existujúcich výskumov reči v laboratóriách by sme ako hovoriaci mali preferovať viazanie gramatických slov viac ako lexikálnych. Keďže tento výskum nie je extenzívny, nemôžeme existujúce výskumy v preferencií vyvrátiť. Otázka viazania gramatických a lexikálnych slov v tejto práci teda ostáva otvorená.

Ďalší výsledok sa týka porovnania jednotlivých hovoriacich medzi sebou v počte viazania a glotalizácie. V tejto časti boli porovnaní dvaja hovoriaci z každej skupiny, jeden muž v porovnaní s jednou ženou. V porovnaní amerických hovoriacich mala žena veľmi nízky počet viazania oproti počtu glotalizácie. Dôvodom bolo veľmi pomalé tempo reči. Naopak, americkí hovoriaci, ktorý bol muž mal zase vysoký podiel viazania oproti glotalizácií. Dôvodom bolo opäť tempo reči, no na rozdiel od predošlej hovoriacej mal tempo rýchle. Porovnanie britských hovoriacich bolo veľmi podobné ako porovnanie amerických hovoriacich. Britská hovoriaca mala najvyrovnanejší počet viazania a glotalizácie zo všetkých v bristkej skupine. Britskí hovoriaci mal naopak veľký extrém v hodnote viazania. Dôvodom mohla byť dynamická reč a rýchle tempo, ale aj dĺžka danej nahrávky. Ako bolo spomenuté vyššie, nahrávky mali rôzne dĺžky a tento hovoriaci mal jednu z najdlhších nahrávok vôbec.

V ďalšom bode bol porovnávaný počet viazania a glotalizácie u individuálnych hovoriacich s dôrazom na slovný druh. Hovoriaca americkej angličtiny mala veľmi nízky počet viazania v gramatickej kategórií ako iní hovoriaci. Počas opakovaného počúvania nahrávky bolo očividné, že hovoriaca má tempo reči skôr pomalé a využíva zdôrazňovanie slov. Traja ďalší hovoriaci, dvaja americkí a jeden britský, mali hodnoty viazania v gramatickej kategórií cez 75 percent. Tu môžeme zase vidieť rýchle tempo reči oproti

ostatným hovoriacim. V lexikálnej kategórii sme sa zamerali na dvoj hovoriacich, jednu americkú hovoriacu a jednu britskú hovoriacu. Americká hovoriaca viazala lexikálne slová málo, celková hodnota bola 25 percent. Britská hovoriaca mala hodnotu viazania cez 90 percent a v gramatickej kategórii bola hodnota viazania tiež pomerne vysoká. Vraciam sa späť ku argumentu, že britskí hovoriaci mali prirodzene rýchlejšie tempo reči, čo môžeme vidieť na týchto výsledkoch.

Posledný výsledok v tejto kapitole sa týka počtu viazania a glotalizácie s ohľadom na samohlásky a spoluhlásky. V tejto časti je potrebné spomenúť, že vo výsledkoch nie je započítavaná glotalizácia samohlások po pauze, viedlo by to k prílišnému ovplyvneniu výsledkov. Ako bolo spomínané vyššie, britskí hovoriaci majú rýchlejšie tempo reči, tým pádom hodnoty glotalizácie a viazania bývajú vyššie oproti americkým hovoriacim. V tejto časti je tento rozdiel vidieť veľmi dobre. Britskí hovoriaci mali počet spoluhlások vo všetkých nahrávkach cez 350, pričom americkí hovoriaci ich mali iba niečo málo cez 200. Viazanie samohlások sa v oboch kategóriách zdá byť vyrovnané. Obe kategórie hovoriacich ale preferujú viazanie spoluhlások viac ako viazanie samohlások.

Kapitola číslo 5. sa venuje všeobecnej diskusii a záveru celej práce. Táto práca ponúka náhľad to problematiky viazania v spontánnej reči a tendenciám britskej a americkej angličtiny v ohľade viazania. Z tejto práce vychádza, že obe angličtiny preferujú viazanie spoluhlásky so samohláskou viac ako viazanie samohlásky so samohláskou. Otázka či môžeme aplikovať existujúce poznatky zo štúdií laboratórnej reči, preferencia viazania gramatických slov viac ako lexikálnych, ostáva otvorená vzhľadom na nedostatok dát. Pre budúci výskum by bolo potrebné zanalyzovať viac hovoriacich aby mohli byť výsledky štatisticky presvedčivé. Ďalší výskum tejto témy by mohol pomôcť študentom angličtiny, ako ich druhého jazyka, pri zlepšovaní sa v oblasti plynulosti reči.