ELTE EÖTVÖS LORÁND UNIVERSITY

DIPLOMAMUNKA
M.A. THESIS

Zöngétlen mássalhangzó kapcsolatok az angol nyelvben: Fonológiai elemzés az aspiráció, asszimiláció, és az átírási konvenciókról

Voiceless obstruent clusters in English: A phonological analysis exploring aspiration, assimilation, and transcription conventions

SUPERVISOR

Dr. Szigetvári Péter

Associate Professor

WRITTEN BY

Wutka Vencel

English Studies M.A.

Theoretical Linguistics Specialization

2025

DIPLOMAMUNKA
M.A. THESIS

WUTKA VENCEL

English Studies M.A.

Theoretical Linguistics Specialization

2025

TABLE OF CONTENTS

Pledge of scholarly honestly viii

Acknowledgment ix

Disclaimer x

Abstract xi

1 Introduction **1**

2 Theoretical background **5**

2.1 A voicing system: Hungarian 5

2.2 An aspirating system: English 9

2.3 Postaspiration and syllabification: Potential problems 17

2.4 Laryngeal assimilation: Potential problems 22

2.5 *Fortis+fortis: A phonotactic constraint in English 27

2.5.1 Yorkshire English 30

2.5.2 Durham English 31

2.6 Transcription conventions 32

2.7 Some conclusions 34

3 Experimental procedures **35**

3.1 Hypothesis 35

3.2 Material and data collection 36

3.3 Procedure 39

3.4 Statistical analysis 41

4 Results **42**

4.1 Segmented data representation ... 42

4.2 Segment duration ... 44

4.2.1 Duration of the preconsonantal vowel ... 44

4.2.2 Duration of the postvocalic stop ... 45

4.2.3 Duration of the intervocalic consonant cluster ... 46

4.2.4 Duration of the vowel and consonant cluster ... 47

4.2.5 Total duration of the occurrence ... 49

4.3 Stop voicing ... 51

5 Discussion **54**

5.1 Interpretation of segment durations ... 54

5.2 Interpretation of stop voicing ... 55

5.3 Overall implication of the experiment ... 58

6 Conclusion **60**

7 References **62**

Appendices **67**

LIST OF FIGURES

2 Theoretical background

Figure 1. RLA in Hungarian 6

Figure 2. RLA in clusters in Hungarian 6

Figure 3. Laryngeal Constraint in Hungarian 8

Figure 4. Laryngeal Constraint and RLA in Hungarian 8

Figure 5. Alleged aspiration spreading in English 24

Figure 6. Yorkshire assimilation across a boundary 30

Figure 7. Yorkshire assimilation within a morpheme 30

Figure 8. Durham assimilation 31

4 Results

Figure 9. Segmentation example on absolute (abs_35_low.textgrid) 42

Figure 10. A segmentation example on rhapsody (rha_13_low.textgrid) 43

Figure 11. A segmentation example on knapsack (kna_20_low.textgrid) 43

Figure 12. Distribution and density of the preconsonantal vowel (in ms) 44

Figure 13. Distribution and density of the postvocalic stop (in ms) 45

Figure 14. Distribution and density of the intervocalic consonant cluster (in ms) 46

Figure 15. Distribution and density of the vowel+consonant cluster (in ms) 47

Figure 16. Ratio of the vowel’s duration to the consonant cluster’s duration 48

Figure 17. Distribution and density of the occurrences (in ms) 49

Figure 18. Ratio of the vowel’s duration to the total occurrence’s duration 50

Figure 19. Duration of voicing in the stop consonant (in ms) 52

Figure 20. Voicing ratio in the stop consonant 52

Figure 21. Overlapping density plot and beehive plot for voicing ratio 53

LIST OF TABLES

2 Theoretical background

Table 1. Laryngeal oppositions in an aspirating versus voicing system 15

Table 2. Positions of post-fricative fortis stops in English 21

Table 3. Standardly-assumed three-way allomorphy of the Z and D suffixes 23

3 Experimental procedures

Table 4. Distribution of accent in the videos 38

Table 5. Distribution of gender in the videos 38

Table 6. Experimental parameters 39

4 Results

Table 7. Duration of the preconsonantal vowel (in ms) 44

Table 8. Duration of the postvocalic stop (in ms) 45

Table 9. Duration of the intervocalic consonant cluster (in ms) 46

Table 10. Duration of the vowel+consonant cluster (in ms) 47

Table 11. Ratio of the vowel's duration to the consonant cluster's duration 48

Table 12. Total duration of the occurrences (in ms) 49

Table 13. Ratio of the vowel's duration to the total occurrence's duration 50

Table 14. Voicing in the stop consonant 51

Table 15. Duration of voicing in the stop consonant (in ms) 51

Table 16. Voicing ratio in the stop consonant 51

LIST OF ABBREVIATIONS

C	**C**onsonant
DE	**D**urham **E**nglish
FF	**f**ortis+**f**ortis
FL	**f**ortis+**l**enis
IGCW	**I**nner **G**erman **C**onsonant **W**eakening
IPA	**I**nternational **P**honetic **A**lphabet
LC	**L**aryngeal **C**onstraint
LF	**l**enis+**f**ortis
LL	**l**enis+**l**enis
PLA	**P**rogressive **L**aryngeal **A**ssimilation
RLA	**R**egressive **L**aryngeal **A**ssimilation
SEFV	**S**outh **E**nglish **F**ricative **V**oicing
VD	**v**oice**d**
VLESS	**v**oice**less**
VOT	**V**oice **O**nset **T**ime
YE	**Y**orkshire **E**nglish

PLEDGE OF SCHOLARLY HONESTLY

By my signature below, I certify that my ELTE M.A. thesis entitled: *Voiceless obstruent clusters in English: A phonological analysis exploring aspiration, assimilation, and transcription conventions* is entirely the result of my own work, and that no degree has previously been conferred upon me for this work. In my thesis I have cited all the sources (printed, electronic, or oral) I have used faithfully and have always indicated their origin. The electronic version of my thesis (in PDF format) is a true representation (identical copy) of the printed version.

If this pledge is found to be false, I realize that I will be subject to penalties up to and including the forfeiture of the degree earned by my thesis.

Date: Oct. 13, 2024 Signed: ……………………………

ACKNOWLEDGMENT

I would like to express my gratitude to my supervisor, Péter Szigetvári, whose expertise and continuous feedback were invaluable in shaping the development of this thesis. His guidance throughout this process has provided both clarity and direction.

I am also grateful to Péter Őri and Attila Starčević, whose insightful courses and class discussions have significantly influenced my research. Attila Starčević's input has been crucial in refining many of the ideas presented in this thesis.

Finally, I would like to thank the Department of English Linguistics at Eötvös Loránd University and a fellow student, Rei Iwai, for fostering an environment of intellectual engagement that has greatly contributed to my academic experience and growth.

DISCLAIMER

Please note that the gender of the speakers referenced in the *Material and Data Collection* subsection (§3.2) has been inferred based on available information and may not accurately reflect the speakers' true gender identity. In cases where the gender was not explicitly stated, assumptions were made for the purpose of the acoustic experiment. These assumptions are not definitive and may not correspond to the speakers' self-identified gender. I acknowledge the limitations of this approach and recommend interpreting the results with this context in mind.

While every effort has been made to ensure the accuracy and completeness of the analysis presented in this thesis, I understand and acknowledge that any errors, mistakes, and inaccuracies are entirely my own.

ABSTRACT

This thesis investigates voiceless obstruent clusters in English, focusing on a recent claim that English disallows adjacent fortis segments, a phenomenon governed by the proposed phonotactic constraint: *fortis+fortis (*FF). Building on this assumption, this thesis draws primarily from linguistic theory and research (G. Kiss 2017; G. Kiss & Szigetvári 2020; Szigetvári 2020), addressing topics such as the linguistic systems of voicing and aspirating languages; grouping of segments (fortis vs. lenis); laryngeal phenomena, features, and oppositions; the potential problems posed by aspiration, syllabification, and laryngeal assimilation, while also noting our transcription conventions. This thesis also explores two regional varieties, Yorkshire and Durham English, which exhibit distinct laryngeal phonology.

To empirically test the phonetic behavior of voiceless obstruent clusters, the words *absolute* and *rhapsody* were chosen. Given that the *FF hypothesis predicts that both words should contain an underlying /bs/ cluster, a series of acoustic measurements was conducted using the software Praat, drawing on various randomly selected recordings from YouGlish, followed by a statistical analysis. The word *absolute* is standardly transcribed with a lenis+fortis cluster as /ǽ**bs**əlut/; however, *rhapsody* is traditionally considered to contain a fortis+fortis cluster as /rǽ**ps**ədɪ/, which may not be so. Therefore, this thesis explores whether *rhapsody* should be (re)analyzed with /bs/ instead of the standardly assumed /ps/, which is otherwise present in morphologically complex words, such as *knapsack* /nǽ**p#s**æk/, the third test item in our experiment.

This thesis, in line with previous analyses, aims to challenge transcription conventions, which may be grounded in established orthographic tradition, since the findings indicate that fortis+fortis (FF) clusters are phonologically unattested in *rhapsody*. A longer stop duration, smaller vowel-to-consonant cluster and vowel-to-occurrence ratios, and a larger proportion and ratio of voicing into the stop consonant provide statistically significant evidence to believe that /bs/ is a better underlying representation of this voiceless obstruent cluster, supporting the recent hypothesis that English does not permit the segmental grouping of monomorphemic fortis+fortis clusters. Thus, this thesis, from a theoretical point of view, also argues for a reevaluation of how such clusters should be represented (either as fortis+lenis or lenis+fortis), while keeping in mind the reasons why pronouncing dictionaries may not follow this (potentially problematic when it comes to surface representation) practice.

***Keywords*:** aspiration, acoustic measurements, English, experiment, *fortis+fortis (*FF), grouping of segments, hypothesis, laryngeal assimilation, laryngeal features, laryngeal oppositions, laryngeal phenomena, linguistic systems, phonetic behavior, phonotactic constraint, regional variety, reevaluation, syllabification, transcription conventions, voiceless obstruent clusters.

1 INTRODUCTION

Languages spoken in various parts of the world often make use of a two-way laryngeal distinction between their obstruents (Honeybone 2005: 323; Őri 2023: 1). True voicing and aspirating languages make up the two very common types of linguistic systems (though, they are not the only ones). What sets these two types of systems apart is the phonetic realization of the laryngeal contrast and the phonological behavior of the two obstruent series.

The voiceless vs. voiced distinction is responsible for the most common laryngeal contrast in languages across the globe (Iverson & Salmons 1995: 370). Languages such as those in the Romance (e.g., French, Italian, and Spanish), Slavic (e.g., Polish and Russian), and Finno-Ugric (e.g., Hungarian) families, for example, make a contrastive use between actively 'voiced' and 'voiceless' obstruents, thus they exhibit a true voice pattern. Accordingly, in these languages, the distinction lies on a series of (pre)voiced and a series of plain voiceless, unaspirated segments (Beckman et. al 2013: 2), in other terms, the presence or absence of active voicing is liable for the two-way laryngeal opposition in this linguistic system.

The second most common laryngeal contrast in linguistic systems, which does not categorize its segments based on voicing, is the group of aspirating languages (Iverson & Salmons 1995: 370). These languages contrast their two series of obstruents based on certain phonetic cues, including (but not exclusively) the presence vs. absence of aspiration, shortening of the preceding sonorant interval, and the strength of articulation (Harris 1994; Iverson & Salmons 1995; Honeybone 2005; Beckman et. al 2013). It is evident at this point that voicing is not a universal characteristic or a feature shared by all languages. For instance, Germanic languages (e.g., English, German, and Icelandic) fall into the category of aspirating languages.

The standard analysis of an English-type of aspirating language often involves the terms: 'voiced' and 'voiceless' for the grouping of segments, which may be misleading given that there is no active voicing of the so-called 'voiced' series of obstruents, which is otherwise the case in a Hungarian-type of true voicing language. As a matter of fact, in an aspirating system, obstruents may only be voiced passively (or perseveratively) in intersonorant position. Unless this criterion is satisfied, the traditionally labeled 'voiced' series of obstruents remains voiceless in English, and as we will see this kind of voicing appears to be phonetic in English.

Laryngeal phenomena and oppositions, similarly to the voicing distinction, are rather system-specific. Dynamic phonological processes, such as laryngeal assimilation in obstruent clusters, particularly of phonation type, is only made possible in a voicing system due to the distinctive feature [voice],[1] which is relevant from a laryngeal point of view, being able to spread backward. As aspirating languages lack such a specification, this would be unexpected and disallowed in a system like that of English, yet it is often standardly assumed that some kind of a 'voicing' assimilation, or Progressive Laryngeal Assimilation (PLA), or aspiration spreading, does occur in English. The standard account maintains that the feature [spread glottis] triggers aspiration spreading to the suffixes: *-s* and *-ed* as well as to the approximants, which, we will see, is, similarly, phonetic rather than phonological in English.

Transcription systems used to visualize the underlying representation of English obstruents often assume, possibly stemming from the spelling, that for the analysis of a two-member obstruent cluster there may be four possible candidates. Traditionally, English is said to allow lenis+lenis (LL), lenis+fortis (LF), fortis+lenis (FL), and fortis+fortis (FF) clusters; however, recent research (G. Kiss 2017; G. Kiss & Szigetvári 2020; Szigetvári 2020) indicate that this may not be so. English appears to disallow two adjacent fortis obstruents, which is the reason why Szigetvári (2020) posits the *FF phonotactic constraint, which eliminates monomorphemic FF clusters by either reanalyzing them as LF or FL, which can easily be visualized with a well-known fact that aspiration is suspended in /s/+C clusters in English.

[1] I assume and use privative/unary features in this thesis. The definition of a privative feature depends on whether a gesture is present or absent. Assuming privative features establishes "a clear relation between the phonetic cue and the phonological feature" (Beckman et. al 1995: 31), and e.g., the negative value of a feature often does not affect phonology, so it is better to assume a single privative feature which may be active in phonological processes.

While the distinction between the two series of obstruents has been subject of both phonetic and phonological research, the laryngeal identity of obstruents in obstruent clusters is relatively underexplored. Since transcription conventions often overshadow the true identity of the underlying segments, this thesis aims to present a phonetically-grounded acoustic analysis of voiceless obstruent clusters in English. After reviewing our theoretical background on linguistic systems, aspiration, laryngeal assimilation, and transcription conventions, this thesis investigates whether the proposed *FF phonotactic constraint, holds in the experiment and analysis presented in this thesis. The triad *absolute*, *rhapsody*, and *knapsack* were selected in order to explore how the obstruents behave in the <abs> sequence of *absolute* and in the <aps> sequences of *rhapsody* and *knapsack*, the former is considered to be a LF, while the latter two are traditionally believed to be FF clusters.

I hypothesize that what is traditionally considered an FF cluster in *rhapsody* may be disregarded as a possible grouping of segments by assuming that *absolute* and *rhapsody* share an identical underlying LF cluster, whereas the morphologically complex *knapsack* has an FF cluster. The alleged existence of FF clusters in monomorphemic words may potentially call for a reevaluation in Standard English, challenging the established view influenced by spelling.

Nevertheless, atypical varieties of English may permit FF clusters such as Yorkshire English, which appears to display distinct laryngeal phenomena with the feature [spread glottis] playing a marked, active role. In contrast, Durham English also seems to be an atypical variety of English, because it has [voice] as an active, specified feature value, though it seems that only an apparent 'spread of voicedness' operates in this variety of English.

In the following section (§2), I provide a detailed description of linguistic systems, providing a theoretical framework to this thesis. This begins with an exploration of the voicing system of Hungarian (§2.1) and continues with the aspirating system of English (§2.2), dealing with laryngeal features, oppositions, and phenomena from a nontraditional perspective, within the framework of laryngeal realism. I then address potential issues related to aspiration and syllabification in §2.3, followed by a discussion of laryngeal assimilation in §2.4. In §2.5, I discuss the newly proposed phonotactic constraint, *FF, and also bring examples of reanalysis. This section also includes an examination of two 'peculiar' varieties: Yorkshire English (§2.5.1) and Durham English (§2.5.2). Transcription conventions are noted in §2.6, leading to some general conclusions presented in §2.7.

In the experimental procedures section (§3), I first present the hypothesis in §3.1, followed by an introduction to the material and data collection methods used for analysis in §3.2. The procedure (§3.3) and the statistical analysis of the experiments (§3.4) are then described, focusing on the investigation into voiceless obstruent clusters in *absolute*, *rhapsody*, and *knapsack*. The results of the experiment is presented in §4, concentrating on segment duration, stop voicing, and various ratios. In §5, I discuss the experiment in relation to the *FF phonotactic constraint, offering theoretical implications. Finally, in §6, I summarize the findings and draw conclusions for the thesis.

2 THEORETICAL BACKGROUND

2.1 A Voicing System: Hungarian

Languages often show diversity in their phonetic and phonological systems concerning obstruent consonants, which are typically found in pairs. These obstruents (just like laryngeal phenomena) can be analyzed in a variety of ways depending on the linguistic system or the theoretical framework since, for instance, "cross-linguistic variation in voicing appears to be diverse" (Cho 1990: 153). Laryngeal realism (Honeybone 2005) is the direction this thesis aims to take, mainly because it appears to adequately capture the underlying representations of these obstruents while also allowing for privative[2] features in the analysis[3] that follows.

Given the fundamental phonological difference in voicing vs. aspirating languages and our privative framework, only the prominent series of obstruents receive a laryngeal feature. In a voiceless vs. voiced language like Hungarian, a two-way laryngeal contrast is based on the phonetic difference of voicing between obstruent consonants, because the production and pronunciation differ as to how the vocal folds are set: laryngeal opening (abduction) produces a voiceless [p] vs. laryngeal narrowing (adduction with medial compression) and vibration produces a fully (pre)voiced [b] (G. Kiss & Szigetvári 2020: 135; Beňuš 2021: 99).

As the voiced obstruent series is responsible for and triggers phonological processes in a Hungarian-type of voicing language, the obstruents /b, d, g, dz, dʒ, ɟ, (v),[4] z, ʒ/ need to be phonologically specified for either the laryngeal feature [voice] (in most traditional accounts), which this thesis subscribes to, or their equivalent alternatives: [slack] (Halle & Stevens 1971), GT 'glottal tension' (Avery & Idsardi 2001), or the |L| in Element Theory (Backley 2011), whereas the voiceless series remains unspecified. Thus, the voiced series has the potential to

[2] Cho (1990: 153) claims that a privative theory is also more desirable due to its general restrictive nature.
[3] I follow the standards set by Kenyon and Knott's (1953) pronouncing dictionary for the transcription of words in this thesis, unless there is a need for different symbols e.g., Durham English or Hungarian.
[4] The Hungarian <v>, especially before a vowel, is a labiodental approximant /ʋ/.

cause the well-known phonological process of Regressive Laryngeal Assimilation (RLA) or spreading to reassociate their feature (Cho 1990: 141). For instance, the Hungarian word ***húszból*** 'from twenty' is realized as [hu:**zb**o:l], due to the above-mentioned process of RLA, where the (pre)voiced /b/ licenses, or spreads regressively, its [voice] feature to the preceding unmarked /s/, creating a [zb] cluster that 'agrees' in voicing (voiced+voiced), see Figure 1.

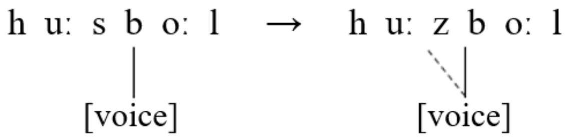

Figure 1. RLA in Hungarian

Similarly, we find the same RLA in obstruent clusters, e.g., in *harasztban* 'in fern' [hɒrɒzdbɒn], where the voiced /b/ reassociates its [voice] feature to the preceding two-member /st/ cluster, creating a [zdb] cluster that 'agrees' in voicing throughout, see Figure 2.

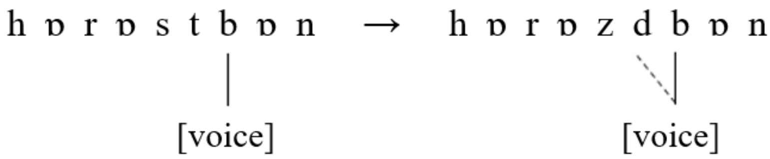

Figure 2. RLA in clusters in Hungarian

The voiced series of obstruents are produced with active voicing in this language, requiring an additional effort by slackening the vocal folds which results in negative Voice Onset Time (VOT). VOT is defined by the timing of the beginning of voicing (usually in the following vowel) relative to the release of a stop closure (Ladefoged & Disner 2012: 139; Beckman et. al 2013: 6). A voicing language then contrasts –VOT (voicing lead) with zero VOT (short-lag), an essential opposition that leads to different realizations between the voiced vs. voiceless obstruental series. Lisker and Abramson (1964) illustrate the average measured VOT between the two stop series in Hungarian: [b] = –90 ms, [d] = –87 ms, [g] = –58 ms, while [p] = 10 ms, [t] = 16 ms, [k] = 29 ms (cited in Beckman et. al 2013: 7). Thus, –VOT means that voice starts during the consonant's closure+hold phase (before release) and continues through the subsequent sonorant interval without pausing (G. Kiss 2021–: §2, ¶28).

Besides that, the plain voiceless series of obstruents /p, t, k, ts, tʃ, c, f, s, ʃ/ in Hungarian contains no laryngeal features in a privative theory given their phonologically inactive nature, and because of the other (oral) features of articulation they remain audible (Honeybone 2005: 324). Even though the voiceless series is not actively involved in phonological processes like feature spreading, in certain voicing languages, such as Hungarian, Czech, Konami, Pengo, Polish, Russian, and Serbo-Croatian, there is still an apparent 'spread of voicelessness.' Since assimilation rules are known to produce clusters with a linked structure, the Underspecification Theory requires [voice][5] to be a universally privative feature, and there is no point in the grammar where [–voice] would be present, the spreading of the negative value of a feature is simply not possible[6] (Cho 1990: 142, 146–147). The apparent spread of voicelessness then requires a different explanation.

Cho (1990: 145) argues that these voicing languages may only allow adjacent voiceless obstruents if the final member is devoiced. Following Mester and Itô (1989), Lombardi (1995: 51) describes how assimilation can be analyzed if [voice] is a universally privative feature and claims that 'assimilation of voicelessness' is simply the result of neutralization because a Laryngeal Constraint (LC) operates in these voicing languages. This constraint involves nonfinal prevoiced obstruents that are not followed by a spontaneously[7] voiced sonorant (vowel and/or sonorant consonant).[8] Their [voice] feature is lost, undergoing delinking, which is the repair mechanism of the LC (Lombardi 1995: 71). Delaryngealization, without the loss of oral specifications, is the end result of this neutralization (Honeybone 2005: 321). The Hungarian word e.g., *háztól* 'from house' clearly shows such a LC, according to which it is

[5] Cho (1990: 153) further argues for privativity in writing that one should assume [voice] to be a privative feature until a language is found in which there is positive evidence, such as geminate integrity effects, indicating that both voiced and voiceless clusters have the same linked structure.
[6] A supposed PLA in Dutch, in which a voiced fricative becomes voiceless following a voiceless consonant, is better treated as devoicing rather than assimilation e.g., *tre*[k] + [v]*ogel* = *tre*[kf]*ogel* (Cho 1990: 143).
[7] This kind of phonation is phonetically the most natural (= the unmarked) way of articulation (Őri 2023: 3).
[8] Sonorants (liquids, nasals, rhotics, semivowels, and vowels) have no specification phonologically as they cannot bear any feature in the lexicon since they remain phonologically inactive and transparent to phonological processes, such as assimilation (Cho: 1990: 146). Consequently, no RLA occurs in e.g., *mésznél* /sn/ ≠ [zn].

realized as [haː**st**oːl]. In this example, the prevoiced /z/ cannot uphold its [voice] feature value and delaryngealizes before /t/ as [s] in the absence of a sonorant segment that would follow, creating a [st] cluster that 'agrees' in voicing (voiceless+voiceless), see Figure 3.

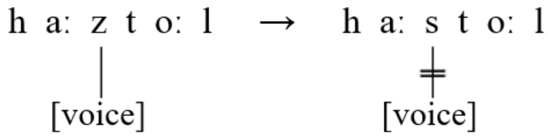

Figure 3. Laryngeal Constraint in Hungarian

Beyond these, Cho (1990: 145) notes that [voice] is lost except from the rightmost obstruent in a cluster, bringing about the interesting fact that we also find delaryngealization of /b/ in words such as *rizsből* 'from rice' [riː**ʒb**øːl], despite a voiced+voiced [ʒb] output. Given that [voice] may only be linked to a nonfinal prevoiced obstruent in the presence of a following sonorant, /ʒ/ cannot uphold its phonation and delaryngealizes to /ʃ/. The LC neutralizes the opposition between the preconsonantal fricative of *rizsből* and *rizstől*, until eventually the /ʃ/ of *rizsből* receives regressive voicing from its voiced neighbor /b/. This neutralization and the subsequent backward spreading of the [voice] is illustrated in Figure 4.

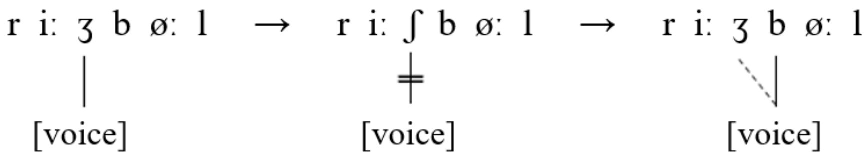

Figure 4. Laryngeal Constraint and RLA in Hungarian

Consequently, we can see that in a Hungarian-type of voicing language, a two-way laryngeal opposition is based on the phonetic difference of active voicing between the obstruent series (G. Kiss & Szigetvári 2020: 135). This leads to the fact that effectively all obstruents in obstruent clusters agree in terms of voicing due to their dynamic phonological behavior (Lombardi 1995: 51). If the rightmost obstruent is voiced, the regressive spread of [voice] produces a cluster of voiced obstruents. If the rightmost obstruent is voiceless, due to the Laryngeal Constraint a voiceless nonfinal obstruent cluster is produced (cf. Figure 4 in which the effects of the Laryngeal Constraint are undone by the regressive spread of [voice]).

2.2 An Aspirating System: English

As we saw, in a Hungarian-type of voicing language, a two-way laryngeal contrast is based on the phonological difference of voicing, as such characterizing segments from the viewpoint of being primarily voiced or voiceless is adequate. Similarly to voicing systems the traditional distinction in an English-type of language involves terms such as ‘voiced’ and ‘voiceless’ (but they are rather system-specific), which may not be applicable in aspirating systems as they fail to accurately capture the phonetic and phonological differences present in these languages.

In an aspirating system there may be, according to Lisker (1986: 5), as many as sixteen acoustic properties that are supposed to cue the identification and difference between the two obstruent series. Pre- and postaspiration,[9] preglottalization, and vowel length differences may be the most well-known and important phonetic cues that influence listeners’ labeling behavior setting the two obstruental series apart from a laryngeal point of view. These distinctions underscore why the traditional [voice]-only analysis is insufficient to account for the opposition in an aspirating system. What is more, these languages lack active voicing (a controlled articulatory movement by either lowering the larynx or fronting the tongue in an effort to delay the devoicing of obstruents), which provides another compelling reason to disregard the standard view, an idea this thesis pursues (Kohler 1984: 163; G. Kiss & Szigetvári 2020: 135).

To add to these, Stahlke (2003: 192) claims that voicing is not crucial to differentiate /b, d, g/ from /p, t, k/. This means that a phonologically ‘voiced’ segment may be phonetically voiceless in an aspirating language (G. Kiss & Szigetvári 2020: 135). Similarly, Abramson and Whalen (2017: 75) argue that a value considered ‘voiced’ in one language may be considered ‘voiceless’ in another. Indeed, based on Jessen (1998), Honeybone (2005: 331) further notes that aspiration is very robust in this linguistic system, and so the underlying contrast cannot be based on voicing. That being so, the feature value [voice], as implied by its broad interpretation

[9] Preaspiration: the glottal opening occurring prior the oral closure (cf. Page 1997).

(Hall 2001: 32), cannot be the sole basis of laryngeal opposition in English. The traditional [voice]-only interpretation does not allow an aspirating system to have a different underlying laryngeal distinction from a voicing system, and maintains that obstruental differences lie in phonological implementation, such as the lack of a voicing rule, that sets them apart and explains their allophonic and output differences, meaning that in this sense Hungarian and English share the same underlying features and phonemes (Honeybone 2005: 329–330).

However, [voice] is not a universal feature that can be used to differentiate obstruents. In an English-type system, [voice] is absent which, for example, stems from the fact that we find no phonologically testable spread of it. Adding to this, [voice] should not be applied to a system whose fundamental laryngeal opposition is different from a voicing one.

Consequently, alternative terms have been proposed to show such a difference between the obstruental series usually available in aspirating languages. I use Sievers' (1876) labels of *fortis* and *lenis* throughout this thesis, as they adequately capture the different nature and various phonetic properties that such obstruents may or may not have. Using a fortis-lenis distinction is useful as it eschews the conventional assumption that the 'voiced' series has a marked feature like [voice], and the "fortis/lenis analysis provides a more general explanation of a wider range of phenomena than the voiced/voiceless analysis" (Stahlke 2003: 191).

This brings us to the point that a language like English has a spread glottis, not a voicing distinction (Abramson & Whalen 2017: 79), an analysis that is extensively argued for by Honeybone (2005), who reanalyzes the Inner German Consonant Weakening (IGCW) and the Southern English Fricative Voicing (SEFV), which contain phonological processes that cannot be accounted for under the traditional [voice]-only analysis. Beckman et. al (2013: 13–14) argue that the traditional approach supposes that these varieties contain a single series of voiced obstruents, a system that does not exist. Under the [spread glottis] analysis both the IGCW and SEFV changes can be analyzed and accounted for. Honeybone's study, for example, supports the idea that the fortis series contains the salient property in an aspirating system, thus specified either as [spread glottis] (Iverson & Salmons 1995; Honeybone 2005), which this thesis subscribes to, or its equivalents: [+fortis] (Kohler 1984), [aspiration] (Lombardi 1994), [+tense] (Jessen 1999), GW the 'glottal width' dimension (Avery & Idsardi 2001), or the |H| in Element Theory (Backley 2011).

The fortis series in English involves the following obstruents: /p, t, k, tʃ, f, s, ʃ, θ/, and even though all of them share the specification [spread glottis], not all of them may be aspirated. Fricatives, unlike stops, lack an audible release phase and have an extremely short VOT, so the presence of [spread glottis] does not always imply an aspirated fricative, though it is a possibility cross-linguistically (Vaux 1998; Abramson & Whalen 2017: 81; Starčević 2024: 3).

The distinction for utterance-initial stops /p, t, k/ and the affricate /tʃ/ is carried by the release phase or the glottal activity (Kohler 1984: 153). An accompanying expulsion or 'puff' of air is a main defining reflex manifestation associated with spread glottis (Iverson & Salmons 1995: 369). In order to produce voicing in the subsequent vowel (or sonorant consonant), Iverson and Salmons (1995: 370) report that the abducted vocal folds start to come together after the release of an oral closure; however, it takes some time for the vocal folds to reach the adducted state required for phonation. This means that the vocal folds are not in contact and the air exiting the trachea during the post-release voicelessness phase is perceived as aspiration.

This aperiodic noise-like sound as a delay in voicing means the VOT is relatively long, leading to the articulation of voiceless aspirated stops. Beckman et. al (2013: 8) illustrate the average values of positive or long-lag VOT in English stop consonants based on an investigation conducted by Lisker and Abramson (1964): [p^h] = 58 ms, [t^h] = 70 ms, [k^h] = 80 ms. The interval that passes between the stop being released and the onset of glottal vibration (or phonation) of the next sonorant is noticeably longer which "carries intuitively a rather well-defined phonetic interpretation" (Nance & Stuart-Smith 2013: 130; Starčević 2024: 3).

Additionally, aspiration appears to correlate with the degree of stress. Keating (1984: 306–308) claims that fortis stops in the beginning of stressed syllables are more heavily aspirated than those beginning unstressed syllables (cited in Iverson & Salmons 1995: 379). Therefore, we find strong (or heavy) aspiration before a stressed vowel initially, e.g., ***cat*** [**k^{h}**æt] and ***pa**per* [**p^{h}**ép$^{(h)}$ɚ], or before a stressed vowel medially, e.g., *po**ta**to* [p$^{(h)}$ə**t^{h}**é**t$^{h\sim(h)}$**o]. We may find weak (or light) aspiration in unstressed positions, e.g., ***po**tato* [**p$^{(h)}$**ət^{h}ét$^{h\sim(h)}$o] and *pa**per*** [p^{h}é**p$^{(h)}$**ɚ], because the glottal opening appears to be smaller in unstressed than in stressed positions, leading to shorter VOT, thus less aspiration (Iverson & Salmons 1995: 376). Though, we may never find aspiration before a strong boundary, as in the final stop of *ca**t*** [k^{h}æ**t**].

Kohler (1984: 154) argues that it is more difficult "to separate [the] presence and absence of aspiration clearly" in utterance-final position or before silence. Therefore, other phonetic cues, such as the strength of articulation, segment duration, prefortis clipping, or preglottalization enable the identification of fortes. Prefortis clipping (Jones 1957; Harris 1994; Wells 2008; Jansen 2007; Cruttenden 2014) is a well-known phonetic phenomenon that may shorten the preceding sonorant interval, leading to a significantly shorter phonetic realization before a fortis obstruent, cf. *c**a**t* [k^{h}**æ̆**t] vs. *c**a**d* [k^{h}**ǣ**d̥]. Preglottalization may accompany the fortis stops /p, t, k/ and the fortis affricate /tʃ/ in final position, another cue only associated with fortis obstruents, found in *ca**t*** [k^{h}æ̆**ˀt**] for example. What we can glean from this is that "the laryngeal contrast is not neutralized…in final position either" (Őri 2023: 4).

Spread glottis results in the variable VOT distributions commonly observed in English. Typically, the singleton fortis stops /p, t, k/ and affricate /tʃ/ exhibit a substantial lag between the release and the onset of voicing. However, in obstruent clusters such as /ps/ or /sp/ (the latter is discussed further in §2.3), this delay, in the absence of a following spontaneously voiced sonorant, may not occur. This pattern parallels the unaspirated singleton stops in Hungarian, which are produced without a delay in VOT (Iverson & Salmons 1995: 374).

When /p, t, k/ is followed by /l, r, j, w/ aspiration manifests itself in the devoicing of the following non-nasal approximant, which otherwise remain predictably voiced after voiceless stops in a true voicing system (Iverson & Salmons 1995: 373). For example, the Hungarian ***pl**ató* [**pl**ɒtoː] shows no devoicing, also capturing why [–voice] cannot be a feature. In contrast, in an aspirating system, the conventional assumption implies that fortes may trigger phonological processes. The supposition of PLA or aspiration spreading is conventionally applied to the English approximants, such as in ***pr**ay* [**pr̥**eɪ], and to the the *-s* and *-ed* suffixes, such as in *pac**ked*** [p^{h}æ**kt**]. Nádasdy (2008: 63) subscribes to a "complete voicing assimilation" which governs the 'voiceless/voiced' nature of the suffixal obstruent. This view is challenged in §2.4 as devoicing seems to be a phonetic, rather than a phonological matter in English.

On top of that, /t/ is the only fortis stop which may be subject to voicing and lose its [spread glottis] feature. Speakers of flapping accents of e.g., American English, may pronounce *a**t**om* and *A**d**am* as homophones with an alveolar flap [ǽ**ɾ**əm] (Iverson & Salmons 1995: 376).

Contrarily, the lenis series of obstruents /b, d, g, dʒ, v, z, ʒ, ð/ shows none of the previously mentioned phonetic cues. These obstruents may never be aspirated, glottalized, and may never shorten the preceding sonorant interval, hence their unmarked state in an aspirating system like English. When these stops are in utterance-initial positions, voicing may begin just before the release of the stop, which means that they are produced with zero VOT, and in many cases they may be completely voiceless: [b̥~p] = 1 ms, [d̥~t] = 5 ms, [g̊~k] = 21 ms (Ladefoged & Disner 2012: 139; Lisker & Abramson 1964, cited in Beckman et. al 2013: 8). Adding to this, in isolated utterances, a lenis /b/ is likely to display a silent interval (though some aperiodic noise may be emitted), followed almost instantly by the onset of voicing, which means that it may resemble a voiceless [p] perceptually (Abramson and Whalen 2017: 79; Beňuš 2021: 125).

The only position where a lenis obstruent may receive voicing is between sonorants. This kind of voicing is different in nature from the intervocalic voicing found in a voice-

voiceless system. While obstruent voicing is phonologically active in a voicing language, it is phonetic in an aspirating one (Backman et. al 2013: 5, 29, 31). The reason why the lenis series may be voiced, though only phonetically, has to do with the assumption that phonetic processes cannot change a specified phonological feature, meaning that passive voicing may only affect unspecified lenes in an intersonorant, voice-friendly position in English (Jessen & Ringen 2002, cited in Beckman et. al 2013: 36).

Perseverative voicing is achieved with the continuation of voicing throughout an intersonorant obstruent interval. This is the only position where voicing is a distinguishing trait between a fortis vs. lenis obstruent in this system. Abramson and Whalen (2017: 79) argue that glottal pulses proceed unbroken from the preceding vowel in lenes, hence Lisker and Abramson's (1964: 417) labels of "unbroken voicing" and "edge vibrations" or Davidson's (2016: 43) "bleed." This means that intervocalic lenis stops have a negative VOT, therefore a voicing lead can be associated with these medial obstruents in English.

Kohler (1984: 162–163) claims that certain aspects of sound production "for its continuation," such as a high degree of adduction, a low degree of vocal fold tension, lax walls, and movable parts, "are diametrically opposed to those necessary in fortis obstruents." These physiological conditions of voicing may only be associated with intersonorant lenes. Unless these conditions are satisfied, lenes remain partially (or even fully) voiceless elsewhere. And even in intersonorant position, Docherty's (1992: 120) reveals that they may only be partially voiced, ½ to ⅔ (~57%) of a lenis stop's closure tends to be voiced, while the rest voiceless.

Beňuš (2021: 119, 125) claims that even though the vocal folds may maintain voicing at least for the initial part of the closure for /b/ (/p/ tends to have a fully voiceless closure), what we may find is a devoiced [b̥] that perceptually sounds similar to [p] due to the phonetic similarity of VOT. As a result, a lenis obstruent has two allophones: a passively voiced one as in *pa**dd**ing* [pʰæ**d**ɪŋ] and a partially/fully voiceless one as in *pa**d*** [pʰæ**d̥/t**] (G. Kiss 2021–: §2,

¶39). This highlights that voicing in English depends on the environment and can only play a limited role in distinguishing the obstruents, which is also corroborated by Ladefoged and Disner (2012: 139) who argue that an English lenis /b/ sounds like a voiceless Spanish /p/.

Overall, both the phonological patterning and the dynamic phonological behavior of the two series of obstruents are generally different from one another in a voicing vs. aspirating language, "support[ing] the idea that the two obstruent series should be represented differently" (Őri 2023: 8). The conventional approach, which claims that "English and other Germanic languages are commonly said to have voiced stops that contrast with voiceless stops," falls short of capturing the underlying featural property of obstruents in aspirating languages (Ladefoged & Disner 2012: 139), calling for a different analysis.

The fortis-lenis analysis, which frames the laryngeal opposition among the obstruents as spread vs. non-spread glottis, not as merely voiceless vs. voiced, offers a better approach since, as we saw, languages may differ as to how the obstruental contrast is preserved.

English-type of aspirating system		***vs.***	**Hungarian-type of voicing system**	
[spread]	/p^h/	<p>	/p/	non-spec
	/t^h/	<t>	/t/	
	/k^h/	<k/c>	/k/	
non-spec	/p/	<b>	/b/	[voice]
	/t/	<d>	/d/	
	/k/	<g>	/g/	

Table 1. Laryngeal oppositions in an aspirating versus voicing system

Table 1 above aims to sum up and illustrate the relationship of spelling and the fundamental phonological reality, considering markedness principles, laryngeal oppositions between the two series of obstruents, and the underlying representation of the two series of stops available in our reference systems. The contrast between English and Hungarian obstruents is best shown in initial position. In English, an orthographic <p> represents a voiceless aspirate [p^h], while <b> is realized as a plain, unaspirated, partially/fully voiceless [b̥~p]. In Hungarian, <p> corresponds to a voiceless unaspirated [p], while <b> is realized as an actively (pre)voiced [b].

Therefore, there is a clear typological difference between the majority of Germanic vs. Romance, Slavic, and Uralic languages. The former contrast partially ~ fully voiceless and passively voiced vs. aspirated stops, while the latter display an opposition between fully voiced vs. unaspirated voiceless stops. The "salient property defining contrast" in aspirating systems is the feature [spread glottis] of the fortis series, and not [voice] which is the feature of the voiced series in voicing systems (Iverson & Salmons 1995: 369). This is known as the narrow interpretation of [voice] (Hall 2001: 32), which in addition to making use of the fortis-lenis distinction, limits the use of the term 'voiced' for elements that are phonologically marked voiced and also reveal "phonologically testable voicing," such as RLA, rather than phonetic, predictable, or passive voicing (Honeybone 2005: 328, cited in Starčević 2024: 3).

Honeybone (2005: 345) refers to this nontraditional method of representing contrast as laryngeal realism, which describes "a set of assumptions rooted in the phonological reality of the laryngeal features as measured against the set of phonological processes likely to result from these" (cited in Starčević 2024: 3). This method accurately captures and explains the laryngeal contrasts in languages with two obstruent series. Therefore, the laryngeal opposition between the examined linguistic systems is best shown by using different subsegmental units, leading to different predictions as to what is a possible phonological behavior. For this reason, Hungarian and English are underlyingly different from a laryngeal standpoint.

Lastly, based on the previously mentioned phonetic cues, it is always possible to determine the laryngeal identity of a singleton obstruent; the laryngeal identity of obstruent clusters, however, is not always detectable (G. Kiss & Szigetvári 2020: 136), raising potential problems regarding aspiration and laryngeal assimilation, detailed in §2.3 and §2.4.

2.3 Postaspiration and Syllabification: Potential Problems

Descriptions (cf. Nádasdy 2008; Stahlke 2003; Trager & Smith 1956, among other analysts) of English phonetics and phonology regularly mention an exception to the basic rules of aspiration: fortis stops are not aspirated in post-/s/ position. Different analyses have been presented to account for this, but some fail to capture the underlying laryngeal identity of obstruents in obstruent clusters and why these stops must be unaspirated lenes.

Some analysts approach the issue with syllabification. According to Wells (1990: 77), the fortis stops /p, t, k/ are aspirated when they are initial in a stressed full-voweled syllable. So, we find an aspirated stop in ***p**ad* [**pʰ**æd̥], ***t**armac* [**tʰ**ármək], or ***c**abinet* [**kʰ**ǽbɪnət]. However, Szigetvári (2020: 45) argues that stops in words such as ***sp**end* [**sp**ɛnd̥], ***sp**aghetti* [**sp**əgɛ́tɪ], or *a**sp**en* [ǽ**sp**ɛn] are not aspirated at all since aspiration appears to be suspended in /s/+stop sequences. Contrary to the traditional assumption, tackling the issue of a deaspirated stop in /s/+stop clusters with syllabification is problematic for many reasons.

The vowel+obstruent cluster /abst/ of *ab**st**émious* [əb**st**émiəs],[10] for example, may be syllabified at least four ways as different syllabification principles generate different results. The first option is to place the syllable boundary between the word-initial vowel and the obstruent cluster, thus *a.bstémious*. Nevertheless, this syllabification may not be allowed. Kaye (1992: 1) argues that according to the Binary Theorem all constituents are maximally binary, and so an onset may not contain more than two consonants in English. Therefore, a three-member onset like /bst/ is ill-formed, and generally no word may begin with it. Sonority sequencing principles also justify that /bst/ is an illicit initial cluster in English. Since syllables display sequences of "quasiperiodic rise and fall in sonority" as part of the sonority cycle, and "onsets have a rising sonority in English," /bst/ would form a rising-falling sonority, which

[10] I only focus on the [əbst] sequence as aspiration-wise this is the important part; however, [m] depending on the principle may be part of the coda of the first syllable (as the outcome of coda capture), or of the onset of the second syllable, or it may even be an ambisyllabic medial consonant as a stressed vowel cannot stand syllable-finally despite Kahn's attempt of onset maximization (Kahn 1976: 47).

may not constitute a well-formed onset (Clements 1990: 299; Törkenczy 2015: 11). Selkirk's (1984: 112) detailed sonority scale assigns a sonority value to every phoneme in English, for example /s/ = 4, /b/ = 1, /p/ = 0.5 (the lowest score). Therefore, Selkirk would subscribe to the prediction that /s/ is too sonorous to stand before /p/,[11] which apparently makes it impossible for sonority to gradually rise before and fall after the nucleus of the syllable, so /bst/ is in violation of her theory which means that this cluster should be rejected.

The second option is to divide the obstruent cluster, thus *ab.stémious*, which is similarly an unsatisfactory attempt at syllabifying the word. Selkirk's principle, for example, rejects /st/ as a possible grouping of segments, though it is still a permissible initial cluster in English.

The theoretical third option would divide the word after the obstruent cluster and before the stressed vowel, thus *abst.émious*. Though, this method is also problematic in the absence of the Onset Maximization Principle. Kahn (1976: 41) advocates the idea that onsets have precedence over codas since onsets generally allow more options (Lombardi 1995: 71); besides in coda position consonants are often susceptible to a loss (a fact present in many languages). Since a well-formed onset should be of maximal length, *abst.émious* fails to abide by this rule in having a consonant-final syllable followed by a vowel-initial one.

The last remaining option is then *abs.témious*[12] in which /t/ finally occupies a stressed initial position and /bs/ satisfies the Sonority Sequencing Principle, although this can only be achieved if the sonority scale is either reduced to a four-term hierarchy (obstruent < nasal < liquid < glide, Clements 1990: 297) or to "the most common subdivisions of segments with respect to sonority," that is to vowels, sonorants, and obstruents (Zec 1995: 86). Accordingly,

---

[11] Nonetheless, English seems to have plenty three-member clusters e.g., *spr-, str-, skr-, spl-, spj-, stj-, skj-, skw-*, and even the marginal cluster *?skl-*, all of which apparently violate both Kaye's (1992: 1) Binary Theorem and Selkirk's (1984: 112) Sonority Sequencing Principle. Different ideas have been forwarded to solve this issue, one analysis assumes an empty/unpronounced CV sequence before /s/, which is now in the coda; another creates a degenerate syllable with /s/ its only element, attaching either to preceding or following syllable (Aoun 1979: 145).

[12] Kaye (1992: 11–12) notes that this way of syllabification also has some important psycholinguistic backing based on the experiments of Treiman et. al (1991: 11) into /s/+C clusters, which concludes that all /s/-clusters are divided when they occur in the middle of words and are treated heterosyllabically.

all segments in a subdivision share the exact sonority score, this way sonority remains level, otherwise, /s/ would contain a higher sonority index than /b/ (cf. Selkirk 1984: 112). As such, the syllable initial /t/ in *abs.témious* would be expected to show aspiration since it is finally in stressed onset position, yet aspiration is missing. Consequently, “it is not possible to produce a meaningful distinction [concerning the presence versus absence of aspiration] based on syllabification,” for this reason aspiration rules “could fall out of […] syllable-based analysis” (Szigetvári 2019: 101–102; Őri 2023: 6), hence the answer lies elsewhere.

Others have proposed non-syllabic phonological models that are phonetically grounded and may better explain the absence of aspiration in /s/+stop clusters. One such attempt was Davidsen-Nielsen’s (1969: 55–56) preceptory experiment[13] into the phonetic justification of post-/s/ fortis stops in British and American English. The experiment consisted of recordings of *spear*, *spat*, *steam*, *sty*, *scold*, and *score* whose voiceless obstruent clusters had to be identified by the thirty-two participants. After removing the initial [s], the words were played back to the participants. The eight American English speakers perceived lenis /b, d, g/ at a rate of 98%, while the twenty-four British English speakers identified lenis stops at a rate of 92% in their respective dialects of recordings. This means that such segments were in fact produced without +VOT and below the 35-ms threshold[14] to be perceived as fortis (G. Kiss 2017: 9). Therefore, the contrast between fortis and lenis stops appears to be neutralized in this position (Davidsen-Nielsen 1969: 55; G. Kiss 2017: 3).

It is well-known that [spread glottis] is the feature responsible for a two-way laryngeal opposition in English. A fortis stop is normally aspirated in stressed initial position, that is “produced with a relatively large glottal opening during the stricture” (Catford 1977: 115). Nevertheless, this feature value is suspended in /s/+stop clusters. Szigetvári (2020: 46) argues

[13] Davidsen-Nielsen’s findings are also supported by the results of two earlier investigations: Lotz et al. (1960) as well as Reeds and Wang (1961).
[14] Even if there is a small amount of aspiration, Abramson and Whalen (2017: 77) claim that it is not sufficient to perceptually mark the stop as aspirated.

that these segments "[do] not contain any instruction for keeping the glottis spread," which means that aspiration is "masked by the closure phase of the cluster," and its absence can be attributed to a narrowed glottis which does not appear to be sufficiently spread to trigger an aspiration effect or a devoicing effect in a subsequent sonorant (Iverson & Salmons 1995: 376).

A narrowed glottis,[15] evidenced by glottography and laryngoscopy experiments, is associated with unaspirated lenis obstruents, which explains the absence of a voicing lag, the most easily detectable feature of an aspirated stop, a reason why post-/s/ stops (/sp, st, sk/) must then be lenis[16] and interpreted as /sb, sd, sg/ (Catford 1977: 114; Szigetvári 2019: 101). This phenomenon is validated by most perception-based research into native speakers' judgments on post-/s/ stops (G. Kiss 2017: 3). Adding to this, voicing may not even determine lenis stops intervocalically,[17] so the voicelessness of a post-/s/ stop "cannot be taken to be a symptom of its fortisness" (G. Kiss & Szigetvári 2020: 137).

As a result, analyzing post-/s/ stops as lenis appears to be the correct approach in Standard English. This is why Twaddell (1935: 31) advocates for an "equally consistent and unambiguous transcription," suggesting that /sp/ would be more accurately represented as /sb/, similarly perhaps to fricative+lenis stop clusters which are reflected in Welsh orthography (and thus transcription), for example *Sbaen* 'Spain' /sbai̯n/ or *ysgol* 'school' /ə́sgɔl/ (Őri 2023: 6).

What is more, the expected aspiration of stops also appears to be neutralized in /f/+stop clusters, making these segments similar to /s/+stop cluster discussed above. Nonetheless, there is a scarcity of words that would fit this search criteria. G. Kiss' (2017: 4) research, based on the CUBE Dictionary (Lindsey & Szigetvári 2013–), which is a free online searchable pronouncing dictionary of Current British English, reports that generally no word begins with

[15] "[It] progressively narrows until achieving voicing in the following vowel" (Iverson & Salmons 1995: 371).
[16] Describing this phenomenon goes back as far as Twaddell (1935: 31) who contends that after /s/ these stops are unaspirated and should better be considered as lenes.
[17] Davidsen-Nielsen's (1969: 58) investigation of voicing, conducted with British English speakers, revealed that even in intervocalic position [b, d, g] were not fully voiced, which means that "voicing cannot be regarded as any constant feature of these stops in English" as they may completely lack vocal fold vibrations (Catford 1977: 112).

/ft/ and most words that contain this cluster have it in word-final or unstressed position where an aspirated stop in /ft/ as [ft^{h}] would not be expected anyway. A valuable example highlighted by the author is *fifteen*, which reveals that /t/ after /f/ behaves similarly to /p, t, k/ after /s/, from this it follows that deaspiration seems to be the preferred choice[18] (G. Kiss 2017: 9, 11).

The fact that fricative+stop clusters may exhibit partial voicing when a sonorant follows after a boundary (fortes never do) also adds to their composition as fricative+lenis stop clusters, for example in *wisp of smoke* [wɪs**b̥**bəsmoˀk] (Stahlke 2003: 202, cited in Starčević 2024: 41). Consequently, Szigetvári (2019: 101; 2020: 46–47) takes a further step by claiming that all fortis stops following fortis fricatives—not just /s/ and /f/—are lenis as */sp, st, sk/ may be excluded by a phonotactic constraint in English. From all this, the restriction follows that an aspirated fortis stop may only stand after a fricative in three positions in English. Table 2 below serves to illustrate the possible positions of post-fricative fortis stops in English.

A boundary separates the fortis fricative+stop cluster	The fortis stop is preceded by a lenis fricative	A boundary is involved and the preceding fricative is lenis
briefcase [b̥rí**f#k**^{h}es]	*Aztec* [ǽ**z̥t**$^{(h)}$ɛk]	*cheesecake* [tʃi**z̥#k**^{h}ek]
discolor [d̥ɪ**s#k**^{h}ʌ́lɚ]	*Robespierre* [rάb**z̥p**$^{h\sim(h)}$iɚ]	*out-of-pocket* [áʊt#ə**v̥#p**^{h}άkɪt]
misplace [mɪ**s#pl̥**és]	*Rosecrans* [ró**z̥kr̥**ʌns]	*stovepipe* [stó**v̥#p**^{h}aɪp]
mistype [mɪ**s#t**^{h}áɪp]	*Gazprom* [gǽ**z̥pr̥**ɑm]	*whizz-kid* [(h)wɪ**z̥#k**$^{h\sim(h)}$ɪd̥]

Table 2. Positions of post-fricative fortis stops in English (Szigetvári 2020: 47)

Overall, syllabic boundaries and the syllabic position of a fortis stop does not define whether such a segment may or may not be aspirated, this is strictly speaking “a lexical property of the plosive” (Szigetvári 2019: 102), which is heavily influenced by the environment. The post-fortis fricative stops /p, t, k/ are thus better analyzed with their lenis counterparts /b, d, g/ despite the orthography and our established transcription conventions.

[18] G. Kiss (2017: 11) also notes that aspiration may be variably available after /f/. If a word is felt by speakers to be morphologically complex, which is the case with /ft/, aspiration may occur after the fricative as in *fif#teen* [ft^{h}] (or *lieutenant* [lɛft^{h}ɛ́nənt], at least based on the Standard Southern British English pronunciation).

2.4 Laryngeal Assimilation: Potential Problems

Laryngeal assimilation in English is another contentious topic which concerns the laryngeal identity of obstruents in clusters. Although English generally lacks assimilation rules, it is replete with obstruents clusters across morphemes (where assimilation rules are expected to work) that either 'agree' in fortisness or lenisness (Szigetvári 2020: 39). Contrarily, Yorkshire English (§2.5.1) and Durham English (§2.5.2) serve as counterexamples whose laryngeal phonology permit such phonological processes, more on which later.

It is an established position that in a voicing language, assimilation, "particularly of phonation type," operates regressively, as such the last obstruent of an obstruent cluster defines the voicing of the entire cluster, leading to 'agreement' in terms of voicing (Catford 1977: 225; Szigetvári 1998: 224). This general expectation is rooted in the phonological reality of the 'voiced' series of obstruents, hence their marked state: [voice] (Starčević 2024: 15). Therefore, it is not surprising to see in a voicing language that many obstruent clusters are either voiced+voiced or voiceless+voiceless.

The situation is more complicated in English. In terms of laryngeal identity, we find some kind of an 'agreement;' marked FF clusters dominate the field (Szigetvári 2020: 43), even though nonspecified clusters would be expected to occur more frequently. As [voice] does not appear to be present in Standard English (Beckman et. al 2013: 14), we cannot find any phonologically attestable regressive spread of it either. So, the answer lies elsewhere.

Szigetvári (2020: 43) counted the number of two-member obstruent clusters in the CUBE Dictionary (Lindsey & Szigetvári 2013–) and reports some curious patterns. First, in word-initial position predominantly FF clusters are found. Second, word-medial FF clusters are ten times more common than LL clusters, a striking imbalance. Third, in word-final position FF clusters similarly outnumber LL clusters. This unexpected prevalence suggests that "the balance tilts strongly towards the marked pole of the contrast, fortis" (Szigetvári 2020: 43).

A reason why we find more FF clusters has to do with a supposed complete PLA or progressive devoicing, which is traditionally assumed to be an active process in English, affecting the suffixes: *-s* and *-ed* (Nádasdy 2008: 63; Ringen & Helgason 2004: 62). The suffix *-s* has several grammatical functions: it marks the regular nominal plural, the verbal third person singular present form, and the genitive, while also serves as the cliticized and contracted 'weak form' of *is* and *has*; for Szigetvári (2020: 44) this is the "Z suffix." The suffix *-ed* is responsible for the past participle and the verbal past tense; for Szigetvári (2020: 44) this is the "D suffix." I adapt Nádasdy's description and Szigetvári's labeling to show that traditionally both the Z and D suffixes are said to have three allomorphs, regardless of their functions.

"Z" suffix				
After sibilants	**/ɪz/**	*buses* /bʌ́s**ɪz**/	*clashes* /klǽʃ**ɪz**/	*buzzes* /bʌ́z**ɪz**/
After 'voiceless' consonants	**/s/**	*clocks* /klɑk**s**/	*baths* /bæθ**s**/	*taps* /tæp**s**/
Any other position	**/z/**	*drills* /drɪl**z**/	*games* /gem**z**/	*Mary's* /mɛrɪ**z**/

"D" suffix				
After /t/ and /d/	**/ɪd/**	*fated* /fét**ɪd**/	*fended* /fɛ́nd**ɪd**/	*padded* /pǽd**ɪd**/
After 'voiceless' consonants	**/t/**	*clocked* /klɑk**t**/	*cussed* /kʌ́s**t**/	*tapped* /tæp**t**/
Any other position	**/d/**	*drilled* /drɪl**d**/	*gamed* /gem**d**/	*married* /mɛrɪ**d**/

Table 3. Standardly-assumed three-way allomorphy of the Z and D suffixes (Nádasdy 2008: 63–64)

Allegedly, a stem-final consonant dictates how the Z and D suffixes behave, exhibiting a straightforward alternation of either being fortis or lenis. These suffixes, which then must contain laryngeally unmarked obstruents, are usually assumed to acquire fortisness through an active feature reassociation of [spread glottis] from a stem-final fortis obstruent, the reason why aspiration spreading or PLA is traditionally believed to occur (Harris 1994: 137, cited in Szigetvári 2020: 44). Nevertheless, aspiration spreading in the Z and D suffixes "would be the only case of laryngeal assimilation in English," Szigetvári (2020: 44) argues. Indeed, Starčević (2024: 18) further claims, it is a "very odd assimilation that does not operate morpheme-internally," which brings us to the fact that in monomorphemic words, we may never find aspiration spreading in either direction. The examples include among various others: *Abkhaz*

/ǽ**bk**ɑz/, *Afghan* /ǽ**fg**æn/, *anecdote* /ǽnɪ**kd**ot/, *Aztec* /ǽ**zt**ɛk/, *Leipzig* /láɪ**pz**ɪg/, *Latvia* /lǽ**tv**iə/, *röntgen* /rə́n**tg**ən/, or *obtain* /ə**bt**én/. What we can already glean from this is the idea advanced by Szigetvári (2020: 44), claiming that no such assimilation may occur in the Z and D suffixes either. For instance, the feature [spread glottis] of /k^h/ in *atta**ck**ed* and *ne**ck**s* does not spread to /d/ and /z/ and 'turn' them into fortis /t/ and /s/, despite the traditional assumptions, according to which the stem-final fortis obstruent licenses its [spread glottis] feature which triggers 'agreement' in terms of fortisness: /k**d**/ > /k**t**/ and /k**z**/ > /k**s**/.

Figure 5. Alleged aspiration spreading in English

Therefore, no phonological process appears to be involved in deriving the voiceless [t] in *attacke**d*** and the voiceless [s] in *neck**s*** as assimilation fails to occur. The voicelessness of these obstruents can be explained with a phonetic fact, a morpheme-final phonetic devoicing or neutralization to [t] and [s] in this phonetically challenging position for a lenis obstruent (Starčević 2024: 18), as voicing may only be upheld in voice-friendly positions in English. Iverson and Salmons (1995: 381) similarly argue that "voiced obstruents may not lie outside of voiceless ones" as part of a tautosyllabic devoicing principle.

Since "a voiceless obstruent in an aspirating language is not necessarily fortis" and given the absence of indispensable phonetic cues (e.g., aspiration, length of the obstruent, strength of articulation) that would encode the Z and D suffixes as fortis, classifying them as containing lenes is then a sound analysis (Szigetvári 2020: 44). Stahlke (2003: 201) also argues that the devoicing of the past tense coronal suffix reflects "the predictable behavior of a lenis stop between voiceless glottal states," further supporting our claim that these suffixes contain lenis obstruents. There is some evidence from the Lindisfarne Gospels, suggesting that Old English scribes may have known that the past tense contained a lenis /d/, e.g., *slēp**d**on* 'slept' or *cēp**d**e* 'kept' are rare spelling examples with <d>, a practice which did not continue as the phonetic spelling with <t> prevailed (cf. Hogg 1992: §7.90). Besides, if aspiration spreading worked progressively, we would expect lenis stops becoming aspirated ones, which constantly fails to occur in Modern English as the result of historical processes (Starčević 2024: 18).

This leads to a reconsideration of the three-way allomorphy of these suffixes, reducing it two alternating allomorphs, making the definition simpler. Szigetvári (2020: 45) proposes that "[t]he D suffix has a vowel-initial allomorph after [t] and [d], and a uniform [d] elsewhere, [and] the Z suffix has a vowel-initial allomorph after a sibilant-final stem, and a uniform [z] elsewhere." Thus, it is unnecessary to assume that an FF obstruent cluster is the outcome of PLA; **/ətǽ**kt**/ and **/nɛ**ks**/ are reanalyzed /ətǽ**kd**/ and /nɛ**kz**/. This analysis does not result in any phonetic difference since the stem-final fortis stop has already phonetically clipped the preceding vowels (G. Kiss & Szigetvári 2020: 139). Besides, in this unfavorable[19] position, a lenis obstruent may leave a similar phonetic imprint as a final unaspirated fortis segment.

Nevertheless, output differences, just like in fricative+stop clusters, may be influenced by a distinguishing boundary. For instance, *Lei**pz**ig*, being a monomorphemic word, may show no phonetic devoicing in the environment of /VpzV/, hence "[pz] is a possible rendition of this cluster" (Starčević 2024: 17), although the word may also show phonetic devoicing since [ps] is another possible pronunciation of the word. Whereas *ta**ps*** does show word-final phonetic devoicing and some morphological conditioning with a distinguishing boundary as #ta**p#s** [ps], rather than being triggered by voicelessness assimilation (Cho 1990: 150–151).

The devoicing of the non-nasal sonorants or approximants: /l, r, j, w/ is also traditionally considered to be the outcome of an alleged active phonological process: aspiration spreading or PLA (cf. Harris 1994; Iverson & Salmons 1995). As long as an approximant comes to follow

[19] Lenes are unstable in unstressed/final position in longer words, leading to less of a chance of recovering phonetic cues, cf. *map* [mæ̆ˀ**p**], *mob* [mā**b̥**] vs. *putrid* [pj̊uʔtrɪ**t**], hence the fully voiceless final stop in the latter example. Similarly to final neutralization in German, where there may be no fortis-lenis distinction, only "saliently devoiced allophones," although not complete for every speaker (Port & O'Dell 1985: 456; Beňuš 2021: 122).

a fortis stop, that would be aspirated if it stood before a stressed vowel,[20] the feature [spread glottis] is said to be shared with an adjacent approximant (Iverson & Salmon 1995: 373, 375). However, having seen that there is no active feature to spread in either direction in English, there is no reason to assume that the spread of aspiration may cause a class-changing process, whereby an unmarked approximant (that would normally have spontaneous voicing) acquires the phonological specification [spread glottis] from the preceding fortis stop. As such, a linked structure or the 'bipositionality' of aspiration between a stop and an approximant is not feasible.

This phenomenon is better explained by the articulatory configuration. When an approximant (or a vowel cf. *tech* [tʰɛk/tɛ̥k]) comes to follow a fortis stop, a noticeable devoicing can be argued to be the byproduct of speech production. This devoicing,[21] which is usually partial, is the manifestation of an open glottis causing +VOT, which means that such approximants may only acquire a 'dimension' phonetically, setting them apart from voiced approximants (Stahlke 2003: 204; Starčević 2024: 4), cf. *claim* [**kl̥**em] vs. *glee* [**g̥l**i], *cray* [**kr̥**e] vs. *gray* [**g̥r**e], *cure* [**kj̥**ɚ] vs. *Gue* [**g̥j**u], and *Queen* [**kw̥**in] vs. *Gwynn* [**g̥w**ɪn], as a delayed VOT is only associated with fortes, we may never find such phonetic devoicing after lenes.[22]

Overall, the absence of any laryngeal assimilation prevents the formation of stem-final FF clusters. Stops and fricatives in the Z and D suffixes seem to contain no laryngeal features encoding their fortisness, and regarding them as lenis (they do still make a voiceless imprint similarly to the fortes) appears to be a sound analysis. Similarly, no active process appears to play a role in approximant devoicing, which is brought about by long-lag VOT, which finally means that aspiration spreading appears not to be phonological, but rather phonetic in English.

[20] This is a compelling reason why aspiration requires stressed vowel support.

[21] For many speakers, approximants are also devoiced when fortis fricatives, especially sibilants, come to precede them, for example found in *slight* [**sl̥**aɪt], *shriek* [**ʃr̥**ik], *sue* [**sj̊**u], or *sweet* [**sw̥**it] (Beňuš 2021: 132).

[22] A peculiar sound change in African-American Vernacular English turned the stressed devoiced initial lenis stops [b̥, d̥, g̊] into fully voiced imploded stops [ɓ, ɗ, ɠ] in various areas of the South Central and Lower Great Plains states (Steward 1973: 48, cited in Stahlke 2003: 208). This positional change only happened where aspirated stops are allowed, which "has the effect of maximizing the phonetic contrast between initial fortis and lenis [stops] that is otherwise reduced by the devoicing of initial lenis obstruents" (Stahlke 2003: 208). The following example is adapted from Stahlke (2003: 208): *That **b**eer ain't no **d**amn **g**ood* [ð̥eəˀt**ɓ**íɹ̠ɚ̃ɪˀtnʌ́w**ɗ**éəm**ɠ**ʊ́əd̥].

2.5 *Fortis+Fortis: A Phonotactic Constraint in English

As we saw not only fricative+stop clusters (§2.3), but also stop+stop clusters at the end of past tense forms (§2.4) cannot form FF clusters. English seems to be a 'peculiar' language from a laryngeal point of view: [voice], which would enable RLA, may be entirely dismissed in the analysis I advocate; and FL clusters found in monomorphemic words show that [spread glottis] cannot possibly imply PLA in English. Thus, the voicelessness of Z and D suffixes and the devoicing of approximants are similarly not expected to be the outcome of aspiration spreading.

Consequently, no laryngeal feature seems to be able to spread in any direction in English. And since no two phonologically marked obstruents, FF, may stand adjacent to each other within a monomorphemic word in English, a phonotactic constraint *fortis-fortis (*FF) was posited on the language by Szigetvári (2020: 47). We then find ourselves in the need of reanalyzing FF clusters as either LF or FL, which can be decided with the help of phonetic cues associated with an aspirated obstruent. What follows is three examples using reanalysis based on Szigetvári's (2020) theory, which we turn to presently.

EXAMPLE 1: *réctor* is better analyzed as having an underlying LF cluster as /gt/ for two reasons. First, the fortisness of /t/ is revealed by stress-placing suffixes,[23] at least by the ones which may combine with the word. These suffixes in our example may include *-ial*, *-ious*, *-ic*[24] and change the stress-pattern of their bases by shifting the primary stress to the stem-final vowel (Nádasdy 2008: 224; Törkenczy 2024: 13–14), leading to the likely explanation as to why native speakers identify such consonants as the /t/ of *réctor* fortis. Thus, when suffixed with the stress-placing *-ial*, the aspirated stop is finally realized on the surface in *rec*[**t**$^{\mathbf{h}}$]*órial*. Aspiration calls for a delayed VOT (a silent interval) following the release burst of a stop (Lisker 1986: 6), which is now an easily observable cue in *rectórial*. Besides, we must find a

[23] Also known as stress-fixing, stress-imposing, pre-stressed, or post-tonic suffixes (Nádasdy 2008: 224).
[24] According to Törkenczy (2024: 13–14), the suffixes *-ial* and *-ic* both require the primary stress to fall on the syllable preceding the suffix. However, the suffix *-(i)ous*, requires stress to fall on the syllable preceding the suffix if it is heavy, but on the second syllable preceding the suffix if it is light.

fortis /t/ in both *réc**t**or* and *rec**t**órial* since “plosives in English do not regularly change from lenis to fortis or vice versa” as they show stability regarding laryngeal identity[25, 26] (G. Kiss & Szigetvári 2020: 139). Second, due to prefortis clipping the phonetic rendition of the stressed vowel of *r**é**ctor* [r**ɛ̆́**$^{(ʔ)}$ktɚ] is predicted[27] to be much shorter than that of, for example, *r**á**gdoll* [r**ǽ̄**gdɑɫ]. Szigetvári (2019: 102) crucially argues that “[i]t seems that all that is necessary for pre-fortis clipping to apply is that a stressed vowel be followed by a fortis consonant before the next vowel, within the same morph.” Consequently, the underlying fortis stop in the LF, that is, /gt/ cluster of *ré**ct**or* still indicates prefortis clipping, despite an intervening lenis stop, resulting in shortening of the stressed vowel. We also expect to find a more rapid stop closure and a longer stop duration as additional phonetic cues of fortisness (Stahlke 2003: 200). While, in the case of the LL, that is, /gd/ cluster of *rá**gd**oll*, no such accompanying phonetic cues may be present as lenis stops do not have a shortening effect, but they do have a slower stop closure, and a shorter stop duration. This intervocalic LL cluster is passively voiced as well. By virtue, there is no reason to assume that an orthographic <k> in *réctor* is a fortis stop since it lacks cues which otherwise encode /t/ to be perceived as fortis.

EXAMPLE 2: *ab**st**émious* (cf. §2.3) is better analyzed as having an underlying FL cluster as /sd/. The voicelessness of the stop does not reveal <t> to be a fortis segment. For articulatory reasons all post-fricative stops are unaspirated as they do not reach a required 35 ms threshold, in terms of their voicing lag (+VOT), to be perceived as fortis. Therefore, [tʰ] is not a possible rendition of this word since the feature value [spread glottis] cannot be associated with the stop,

---

[25] This is why no fortis-lenis synchronic alternations may be claimed in Modern English *tooth–teethe* or *wife–wives*, which are all lexicalized by now. Interestingly, the lenis stop /d/ of the Latinate stems *-cede/-ceed* may palatalize to fortis (and not lenis) fricatives: *acce**d**e–acce**ss**ion, conce**d**e–conce**ss**ion, rece**d**e–rece**ss**ion, sece**d**e–sece**ss**ion, succee**d**–succe**ss**ion* (but *divi**d**e–divi**s**ion* or *diver**t**–diver**s**ion*, where a fortis /t/ > a lenis /ʒ/).

[26] Despite this, laryngeal gestures may not have the expected outcome if their articulatory, aerodynamic, or acoustic conditions are not met (Goldstein 2022b: 2–3). Kohler (1984: 156) claims that if the contact between the articulators is shortened and weakened, fortis stops may become lenis and even receive passive voicing: [p] > [b̥] > [b], while lenis stops may turn into voiced frictionless continuants, which may even be deleted: [b] > [β] > [ ].

[27] Similarly, we expect the shortening of the stressed vowel+n sequence in *bent* [b**ɛ̆n̆**$^{(ʔ)}$t], but not in *bend* [b**ɛ̄n̄**d̥] as the vowel is followed by a lenis obstruent. Thus, the same difference is expected in *rector* vs. *ragdoll*.

which must then mean it is an underlying lenis /d/. Additional phonetic cues, such as the slower closure interval and a shorter duration of the stop, similarly help identify the underlying phonological reality of this stop consonant. This time, the postvocalic fortis /s/ may not cause prefortis clipping since it "affects stressed vowels only" (Szigetvári 2019: 102). Nevertheless, the word contains a three-member obstruent cluster, which may only be reanalyzed as containing an underlying /bsd/ sequence due to the phonotactic constraint: *FF.

EXAMPLE 3: *rhá**ps**ody* and *á**bs**olute*, which are the target items of the acoustic experiment detailed in §3f. Based on our theoretical preliminaries, these monomorphemic words may be reanalyzed with an identical underlying LF, that is, a /bs/ cluster as the orthographic <p> in *rhá**p**sody* may lack indispensable cues for fortisness for a variety of reasons. The stop is not expected to be realized as an aspirate in the absence of a voicing lag, especially so since it is followed by another obstruent. What is more, a shorter stop duration, a slower and somewhat voiced stop-closure may similarly signal the perception of this stop as a lenis /b/, contrarily to what the spelling and standard transcriptions suggest. Nonetheless, the stressed vowel is still expected to undergo some phonetic shortening, even though it is not immediately followed by a fortis obstruent (Szigetvári 2019: 102). Therefore, the supposition also maintains that *rhá**ps**ody* [r**æ̆́**(ʔ)b̥**s**ədi] is similar to *á**bs**olute* [**æ̆́**(ʔ)b̥**s**əlut] in having approximately identical (clipped) vowel durations due to the fortis /s/.

EXAMPLE 4: *kná**ps**ack*, which is the third item under acoustic investigation in §3f. This word is expected (similarly to *absolute* and *rhapsody*) to show prefortis clipping, if not more so, due to the vowel's immediate proximity to the fortis stop, thus [n**æ̆́**(ʔ)**ps**æk]. Despite its identical spelling <aps> with *rh**á**psody*, this item should not necessitate reanalysis since its underlying FF, that is, /ps/ cluster stems from its morphological complexity /p#s/, which allows for two adjacent fortis obstruents in English. This (unaspirated) fortis stop is also expected to show a longer stop duration, just like a more rapid and a fully voiceless stop-closure interval.

2.5.1 Yorkshire English

As we saw, Standard English accents generally prevent aspiration spreading in either direction, but interestingly in West- and South Yorkshire, Yorkshire English (YE) is described to be an exceptional variety that allows such a phenomenon (Wells 1982: 366).

YE is an atypical accent of English as it features asymmetrical laryngeal assimilation. This variety appears to have [spread glottis] as an active feature, which means that the fortes trigger phonological processes, manifesting exclusively in RLA. Yorkshire assimilation then unexpectedly creates FF clusters, which are otherwise claimed not to be possible in Standard English due to *FF. Since fortisness triggers assimilation in YE, what we see is a phonologically complete devoicing of the former lenis obstruent when it comes into contract with a fortis one. Wells (1982: 366–377) reports that this process occurs in compounds or across a true word boundary as in *ma**d**#**h**atter* [mǽ**th**ætɚ], for example.

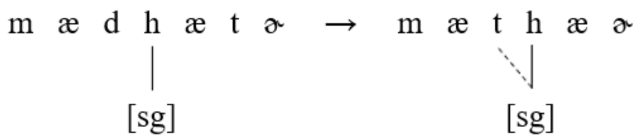

Figure 6. Yorkshire assimilation across a boundary

However, this phonological rule in YE also seems to be true for monomorphemic words so *vo**dk**a* may be realized with a FF cluster as [vɑ́**tk**ə]. This process is in fact not allophonic devoicing but a “complete neutralization of the [fortis-lenis] opposition” (Wells 1982: 377).

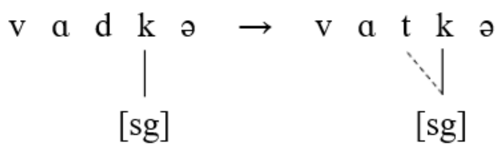

Figure 7. Yorkshire assimilation within a morpheme

What is more, the Yorkshire assimilated *vodka* /tk/ may in fact be glottaled as [vɑ́**ʔ**kə], which provides evidence for a complete assimilation to fortisness. Thus, a three-way allophony of a lenis /d/ is present in YE, which may positionally be a LL [dk], a FF [tk], or even a [ʔk].

2.5.2 Durham English

The sociolinguistic variety known as Durham English (DE) is similarly interesting because its laryngeal phonology unexpectedly appears to contradict Standard English (Kerswill 1987: 42).

DE appears to have a series of prevoiced, unaspirated obstruents corresponding to a series of voiceless, neutral, unaspirated ones (Cyran 2014: 201). This leads Harris (1994: 137) to propose DE to be a voicing system from a laryngeal point of view (cited in Cyran 2014: 201). Therefore, it is fair to assume that aspiration as an active feature is phonologically absent from the system, while active voicing is chiefly present. This is further supported by the phonological behavior of the voiced series of obstruents since RLA occurs across boundaries in the "low-status Durham Vernacular" (which is historically associated with shipbuilders and coal miners); in contrast, RLA is absent in middle-class Durham speech (Kerswill 1987: 42).

Cyran, with *top gun*, illustrates Durham assimilation, in which the laryngeal identity of the rightmost obstruent is passed onto the preceding one, thereby triggering 'agreement.' The actively (pre)voiced [g] licenses and reassociates its [voice] feature to the preceding [p], creating a LL, that is, [bg] (voiced+voiced) obstruent cluster.

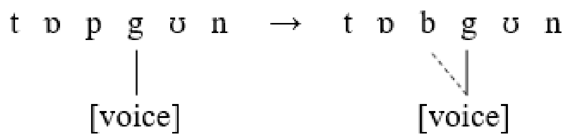

Figure 8. Durham assimilation (Cyran 2014: 202, 204)

At this point, one also expects the apparent 'spread of voicelessness,' which could create voiceless obstruent clusters. However, RLA in DE only occurs if the last obstruent, in a series of obstruents, is voiced. Since this variety also appears to display asymmetrical laryngeal assimilation, a Laryngeal Constraint similar to a Hungarian-type of language (recall *háztól*) is missing from this system. Thus, delaryngealization by delinking the feature [voice], is not a possible outcome. For instance, *big city* [bí**gs**ɪtɪ] preserves its original voiced+voiceless cluster as the prevoiced [g] retains [voice] despite being followed by a voiceless obstruent.

2.6 Transcription Conventions

Standard transcriptions, following the practice of the International Phonetic Alphabet (IPA 2015), are often phonological in nature, do not represent all the information, and are made for pedagogical reasons based on model accents (but they may vary[28] in symbols and conventions).

There is a need for transcriptions to convey pronunciation since languages may have different orthographic systems leading to different degrees of correspondence between spelling and the fundamental phonological reality (Beňuš 2021: 24). This discrepancy may further be burdened by literacy, tradition, cultural background, and orthographic conservatism, which makes spelling "always an imperfect reflection of sound shape" (Starčević 2024: 3; Törkenczy 2015: 2). So, the primary objectives of such transcriptions is to ease these factors, striving for simplicity, greater clarity, and better learnability. The *CMU Pronouncing Dictionary* (Carnegie Mellon University 2014), the *Pronouncing Dictionary of American English* (Kenyon & Knott 1953), the *English Pronouncing Dictionary* (Jones 1917), along the *Longman Pronunciation Dictionary* (Wells 2008) and numerous others are designed with these qualities in mind.

Since a segment may count 'voiced' in one language but may be 'voiceless' in another (Abramson & Whalen 2017: 75), spelling alone may be enough to confuse and mislead, especially, learners of voicing languages. Given the phonological behavior of obstruents in voicing languages, one might completely misinterpret the pronunciation by making a contrastive use of stop categories that consistently differ in voicing (Lisker 1986: 5).

The phonotactic constraint *FF carries significant implications, a possible reanalysis of obstruent clusters is one of many, which may further impede learners' language-learning progress. For instance, if we were to reanalyze and transcribe *spin* based on *FF as [**sb**ɪn], this could mislead speakers of voicing languages, making them believe that the segmental buildup of [sb] could imply RLA as [**zb**ɪn]. Similarly, *cooked* transcribed as [ku**kd**] may encourage

[28] Smith (2000: 293) argues that variation is a good thing, which shows the different ways IPA can be used.

learners of voicing languages to pronounce the word as [ku**gd**]. As a result, by not transcribing words true[29] to their underlying segmental buildup can help learners produce the sounds closer to the native-like pronunciation, without being misled by the *FF phonotactic constraint.

In the same vein, pronouncing dictionaries and beginner textbooks are made for readers with minimal linguistic knowledge and do not wish to go into detail about complexities involving the underlying representation of segments or the use of diacritics representing allophonic variation.[30] Though, in certain cases it would be a very useful (but unlikely) practice of dictionaries to adopt and mark a devoiced, neutral, lenis stop such as [b̥~p] instead of [b], which implies [voice] to be distinctly present, since the linguistic irrelevance of voice difference in English would certainly justify this step (Lisker 1986: 5). This way a dictionary could show that this obstruent has no specified laryngeal features, so it should not invoke a prevoiced [b] for speakers of voicing languages, when in fact it may be a fully voiceless [p]. These segments are phonetically variable in nature and cued by different acoustic properties in different contexts (Lisker 1986: 4). But it may prove difficult for learners to deduce the actual pronunciation from too narrow of a transcription (Smith 2000: 294).

Thus, Lisker (1986: 4) notes that we find no clear indication of phonological preference for transcribing a voiceless lenis /b/ as [b̥] or even [p]. This segment is usually perceived as a partially voiced [b̥] ~ fully voiceless [p], unless it is found in a voice-friendly position which allows perseverative voicing to apply, thus the articulation of a 'fully' voiced [b]. The fact that lenes may be perseveratively voiced, coupled with the aim of trying to be phonologically-minded (rather than being closely phonetic), standard transcriptions mainly stick to [b, d, g] for the representation of initial voiceless lenis stops, for example. This inadvertently contributes to the misanalysis of not only singleton obstruents but also obstruent clusters in English.

---

[29] This is the reason why learners' dictionaries and, in fact, all dictionaries transcribe words the way they do. And, given the recent theory on *FF, not a single dictionary exists that is compiled with this constraint in mind.
[30] Trager and Smith (1951: 30–34) notably propose the largest number of allophones for the English /b/ [b, ᵇ̥b, bᵇ̥] and /p/ [pʰ, p, P, p̚], but it is generally irrelevant what allophones of a phoneme are identified for most dictionaries.

2.7 Some Conclusions

We saw that despite tradition, spelling, and transcription conventions, the segmental buildup of obstruents in obstruent clusters are not as straightforward as it is standardly considered. The dominance of fortis clusters over lenis was suspicious for several reasons, requiring a closer look into (post)aspiration and laryngeal assimilation in English.

While some analysts suggest a correlation between syllable structure and the presence or absence of aspiration, it is clear that syllabification alone cannot fully account for the suspicious loss of aspiration after [s] and [f], which points for alternative explanations beyond the syllable. Experimental evidence highlights the neutralization of laryngeal contrast between fortis and lenis obstruents in post-/s/ positions. Therefore, due to the absence of laryngeal features, which are indispensable cues for fortisness, it is best to assume that in post-fricative positions all obstruents (cf. the three exceptions in Table 2) are lenis for articulatory reasons, despite what orthography and transcription conventions may otherwise suggest.

The long-held view that PLA creates FF clusters in the case of the Z and D suffixes is, similarly, not feasible for articulatory reasons. And on top of that, this would be the only place where assimilation occurs in English, making it a very odd distribution, which does not operate morpheme-internally. On that account, analyzing the suffixal obstruents as lenis appears to be the correct way as they similarly lack any (audible) phonetic cues that could lead to their identification as fortis.

From all this follows that fact that neither fricative+stop, nor stop+stop clusters may be fortis and stand adjacent to each other within a morpheme in English, which leads to the *FF phonotactic constraint posited on the language by Szigetvári (2020). However, in some varieties such as YE such clusters are allowed due to other laryngeal phenomena. Furthermore, as no feature seems to spread in any direction in English, the apparent ‘voicelessness’ of post-fortis segments can be now explained by a phonetic fact, a simple devoicing of obstruents.

3 EXPERIMENTAL PROCEDURES

3.1 Hypothesis

Standard transcriptions, possibly stemming from orthography, often straightforwardly assume the underlying representation of obstruents in obstruent cluster. The items *rhá**ps**ody* /rǽ**ps**ədɪ/ and *kná**ps**ack* /nǽ**ps**æk/ are standardly transcribed with FF clusters, contrarily to *á**bs**olute* /ǽ**bs**əlut/, which is traditionally considered to contain a LF cluster.

The hypothesis, according to which monomorphemic FF clusters are disregarded in English, challenges the standard understanding by predicting that *á**bs**olute* and *rhá**ps**ody* share an identical underlying LF cluster. Our as-if assumption implies that if these clusters behave the same way phonetically, then they are the same phonologically as well, thus /bs/ in both words, irrespective of standard analyses and transcription conventions.

The morphologically complex *kná**ps**ack*, with an intervening boundary between the FF cluster /p#s/, is well-formed according to the hypothesis. As such, /nǽ**p#s**æk/ not only contains a permissible FF cluster because of an intervening boundary (F#F), but it also agrees with the standardly assumed underlying phonological reality of the word.

The hypothesis also predicts that *kná**ps**ack* is the odd one out (and not *á**bs**olute* as the spelling suggests) in having a FF, rather than a LF cluster otherwise present in *á**bs**olute* and *rhá**ps**ody*.

Therefore, I test the robustness of the *FF hypothesis and determine if the underlying phonological reality of *á**bs**olute* and *rhá**ps**ody* is in fact the same and if both words contain an underlying LF, that is, /bs/ cluster (and not FF, that is, /ps/) by making a comparison to *kná**ps**ack*, potentially revealing significant differences and justifying (or not) the hypothesis.

3.2 Material and Data Collection

The experiment presented in this thesis revolved around *absolute*, *rhapsody*, and *knapsack.* The data was supplied by the website YouGlish (https://youglish.com), which utilizes an embedded search engine for YouTube and provides thousands of speech samples of any word or phrase from its database. For the three target items, a list of 50 randomly selected videos was produced respectively, which means that the total pool of videos was capped at a maximum of 150.

After randomly selecting the three sets of videos, I saved the time of the occurrence and the YouTube links on an Excel spreadsheet (Microsoft Corporation 2024) (Appendix A). The data collection was not strictly bound to one reference accent as all seven accents YouGlish provides: "US," "UK," "AUS," "CAN," "IE," "SCO," and "NZ" were selected, albeit "IE," "SCO," and "NZ" accents were arbitrarily not featured. At the time of my investigation, there were overall 30,766 results for *absolute*, 303 for *rhapsody*, and 325 for *knapsack*.

Some videos had to be rejected for a number of reasons. YouGlish mistakenly labeled the portion of the video as containing the target word (2 videos). The speaker's speech rate was too fast, or multiple speakers talked simultaneously (3 videos). The video was deemed to be of overall poor quality (6 videos). Since only one recording was kept for one speaker, multiple entries were rejected (8 videos). The main reason was music during the audio (20 videos). Overall, 37 videos had to be removed (though, some of them had overlaps in terms of rejection, e.g., speakers were talking simultaneously and music was present in the video). After removing the inadequate videos (which were deemed to contain unmeasurable segments), there were 36 usable tokens left for the acoustic analysis of *rhapsody*, 39 for *absolute,* and 38 for *knapsack*.

We assume that the phonotactic constraint *FF is expected to hold for any variety of English which does not display 'peculiar' laryngeal phenomena, cf. YE (§2.5.1) and DE (§2.5.2), as a result voiceless obstruent clusters should be expected to show similar results in general. The gender distribution in the videos was unbalanced with approximately 1.6 times more male speakers. In spite of an approximate 8:5 gender disbalance, this should not impede our measurements as no difference is expected between the speakers' gender this time. See the data distribution broken down by accent and gender at the end of this section (p. 38).

After establishing the base of my experiment, clips were recorded for every remaining item on the list (making sure that there was material before and after the investigated sequence), at a sampling rate of 44,100 Hz, mono while using the internal high-definition audio controller (with the latest drivers installed) of a Gigabyte Z690 UD DDR4 (rev. 1.x) motherboard. These audio recordings were then extracted and saved as uncompressed .wav files, which underwent low-pass filtering with a cut-off rate at 11,025 Hz (to reduce unwanted frequencies) in the open-source digital audio editor software Audacity (ver. 3.6.1, Muse Group 2024), before being saved as .wav files once again. Finally, these were the files used for our analysis in Praat (ver. 6.3.18, Boersma & Weenink 2023). Additional noise removal was not applied since it may also remove certain parts of the speech data, making the audio less natural, besides that the remaining recordings were generally deemed good quality.

The function of the sentence (declarative: 110 tokens; interrogative: 3 tokens) and the sentential position of the target items were not expected to have any significance on the experiment at all since most tokens were located in non-prepausal positions, where boundary-adjacent lengthening[31] was not conditional to occurring. What is more, in the English stress system, longer segments are generally found in stressed syllables compared to unstressed ones (Kim & Cole 2005: 2367), which potentially causes a length difference between stressed (full) and unstressed (reduced) vowels as well, hence the target items were chosen with the same preconsonantal vowel prominence. On top of that, Kim and Cole (2005: 2365) argue that American English tends to exhibit polysyllabic shortening as the temporal consequence of foot-

[31] According to Goldstein (2022a: 5), segments tend to be longer at the end of a phrase, which is the mechanical consequence of approaching a pause.

structure on rhythm, which means that the duration of the stressed syllable is reduced as the number of syllables increases in polysyllabic feet. Even though the target items were selected with the same general criteria in mind, such suprasegmentals as prosodic phrasing and rhythm along with intonation were not considered[32] and thus controlled for in our acoustic experiment.

Lastly, the reason why *absolute* was chosen instead of *absolutely* has to do with a possible frequency effect. The overuse of *absolutely* (173,708 results on YouGlish) in everyday speech may alter the realization of the word, potentially leading to shorter[33] segment durations.

Item	Accent	n	∑
rhapsody	US	28	
rhapsody	UK	5	
rhapsody	AUS	3	36
absolute	US	26	
absolute	UK	11	
absolute	AUS	2	39
knapsack	US	32	
knapsack	UK	4	
knapsack	CAN	2	38

Table 4. Distribution of accent in the videos

Gender	Accent	n	∑
male	US	58	
male	UK	9	
male	AUS	2	69
female	US	28	
female	UK	11	
female	AUS	3	
female	CAN	2	44

Table 5. Distribution of gender in the videos

[32] Kingston and Diehl (1994) note that pitch lowering in neighboring vowels is connected with voicing in obstruents (cited in Iverson & Salmons 1995: 384). 'Pitch lowering' (the decrease of vertical tension by lowering the entire larynx) and 'vocal fry' (the increase of longitudinal tension) were outside of the scope of this thesis.

[33] This effect does not necessarily have to do with faster speech.

3.3 Procedure

The three main parameters measured and calculated (where applicable) can be grouped under segment duration, stop voicing, and ratios (Appendix B). These parameters were necessary in order to determine if the investigated voiceless obstruent cluster in fact contained a fortis or lenis stop consonant in our three target items: *absolute*, *rhapsody*, and *knapsack*.

1) Segment duration
i. Duration of the preconsonantal vowel ii. Duration of the postvocalic stop iii. Duration of the intervocalic consonant cluster iv. Duration of the vowel+consonant cluster v. Duration of the occurrence
2) Stop voicing
i. Voicing present in the stop: "VD" or "VLESS" ii. Duration of voicing in the stop
3) Ratio
i. Ratio of the vowel's duration to the consonant cluster's duration ii. Ratio of the vowel's duration to the occurrence's total duration iii. Ratio of voicing in the stop

Table 6. Experimental parameters

In order to measure segment duration, the sound signal was first divided into temporal stretches (Appendix C). I made use of Praat's annotation window where I manually cut each recording into intervals and assigned labels on a tier (titled: SD) using TextGrid. The onset of the preconsonantal vowel was marked when the formants in the spectrogram and the regular periodicity (vocal fold pulses) in the acoustic waveform became visible. The vowel's end boundary was marked in line with its end formants as the repeating pattern of voicing may continue well into the stop's closure interval (more on which later). The onset of the postvocalic stop coincided with the end point of the vowel. The end boundary of the preconsonantal stop was indicated by a small release noise, if there was any, which occasionally showed an abrupt discontinuity in the acoustic signal. The onset of the fricative roughly concurred with the end of the stop, where noisy waves appeared, while its end point was essentially the end of friction. Since only single 'voiceless' stops are expected to have a voicing lag, VOT measurements were not part of this investigation. Segmentation thus made it possible to measure our experimental parameters: the duration of the preconsonantal vowel, the postvocalic stop, the consonant cluster between the vowels, and the combination of the vowel and the consonant cluster.

Stop voicing was the second major area of investigation in the experiment. Since I did not have access to physiological data on vocal fold vibration, I estimated from quasi-periodic vibrations in the acoustic waveform, and if some background noise was still present (inevitable with some recordings), I was judging by the overall noise level and the spectrogram's readings. If the left edge (or possibly farther) of the stop closure interval displayed some periodic low-amplitude peaks and valleys in the waveform, vocal fold vibrations were detected, which meant that a (somewhat) voiced stop was produced. Sometimes there was an apparent 'coarticulation' from the preceding vowel, and other times there was not. If the waveform was 'relatively' flat (though, there may still be some noise) and there was a lack of periodicity in the stop-closure interval, the stop was deemed voiceless. These items were respectively labeled "VD" and "VLESS," and where applicable the duration of stop voicing was measured, while also keeping in mind when voicing ceases in the stop, on a separate tier (titled: voicing) in Praat.

Calculating ratios to further support the findings proved to be the third important factor. The ratio of the vowel's duration to the consonant cluster's duration was calculated, along with the vowel's duration to the occurrence's total duration. The voicing ratio in the postvocalic stop was similarly calculated, which enabled us to identify whether the 'voiceless' stop was in fact a fortis or lenis obstruent consonant. Lastly, the example segmentations were prepared in and exported from Praat Picture as high-resolution 600 dpi .png files (see pp. 42–43).

3.4 Statistical Analysis

To test whether the hypothesis (*FF) holds across our target items and whether there are differences that necessitate a reevaluation of the conventional phonological classifications of these clusters, standard deviation (SD), standard error (SE), mean, median, min., and max. correlations were calculated in Microsoft Excel (Microsoft Corporation 2024) (Appendix B).

In order to rigorously evaluate the similarities and differences across the target groups a robust statistical analysis was conducted whose objective was to normalize the data by determining the degree of variation between each of the experimental parameters in the three investigated test groups: *absolute*, *rhapsody*, and *knapsack*.

Considering the non-normal distribution of the data and the unequal sample sizes among the groups (39–36–38), I decided to use non-parametric statistical methods to ensure accuracy and reliable results. The Kruskal-Wallis H-test was carried out to test the means across the three groups, with a p-value less than 0.05 indicating if there were significant differences. For the Kruskal-Wallis test, η^2 (eta-squared) was also calculated to measure the effect size and to assess the proportion of group variance. This metric shows how much of the variation in the data is due to the differences between the groups.

Following this, pairwise comparisons with U-statistics were calculated using the rank-biserial correlation from the Mann-Whitney U-test to identify the effect size (r) between the three target groups, which was important for interpreting the size of the differences, providing insight beyond statistical significance alone.

To efficiently process the data and ensure that the correct methods were applied for accurate and reproducible results, two Python 3 scripts were written with the help of ChatGPT (OpenAI 2024) and executed in Google Colaboratory (2024). The statistical script (Appendix D) utilized NumPy and SciPy packages; the Raincloud (Allen et. al 2019) and density plots used for data visualization used Pandas, Matplotlib, Seaborn, and PtitPrince (Appendix E).

4 RESULTS

4.1 Segmented Data Representation

The following figures were prepared as examples for visual representation. These examples were selected based on their representativeness (or prototypicality) from each of the three target groups, approximately being able to capture their average phonetic tendencies.

Figure 9 below shows a typical segmentation of *absolute*. The preconsonantal vowel has a duration of 129 ms, followed by a 61 ms postvocalic stop. The consonant cluster duration between the vowels is 163 ms, the combination of the vowel and consonant cluster is 292 ms, with a total duration of 465 ms for the entire occurrence. The voicing detected in the stop segment lasts 26 ms, thus marked "VD." The ratios calculated from these measurements include a vowel-to-consonant cluster duration ratio of 0.79, an occurrence-to-vowel duration ratio of 0.28, and a voicing ratio in the stop of 0.43.

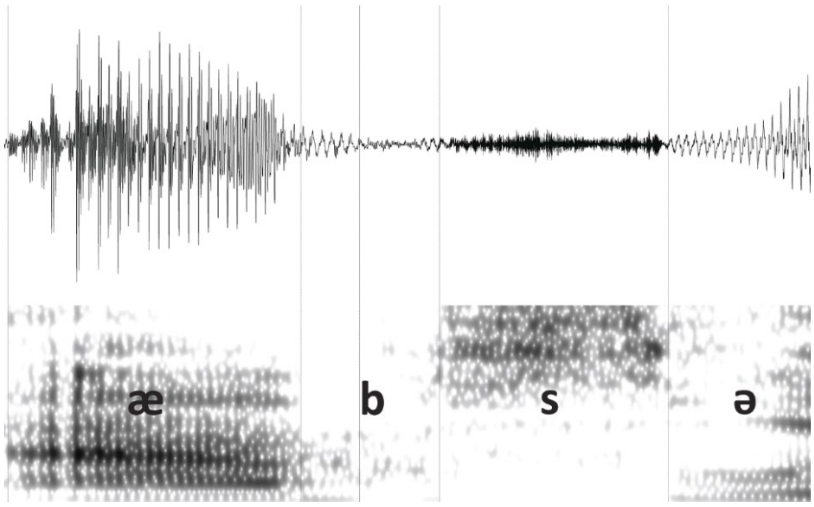

Figure 9. Segmentation example on absolute (abs_35_low.textgrid)

Figure 10 below shows a typical segmentation of *rhapsody*. The preconsonantal vowel has a duration of 135 ms, followed by a 38 ms postvocalic stop. The duration of the inter-vocalic consonant cluster is 153 ms, whereas the combination of the vowel and the consonant cluster is 153 ms, with a total duration of 461 ms for the entire occurrence. The voicing detected in the stop lasts 16 ms, thus marked "VD." The ratios calculated from these measurements include a vowel-to-consonant cluster duration ratio of 0.88, an occurrence-to-vowel duration ratio of 0.29, and a voicing ratio in the stop of 0.42.

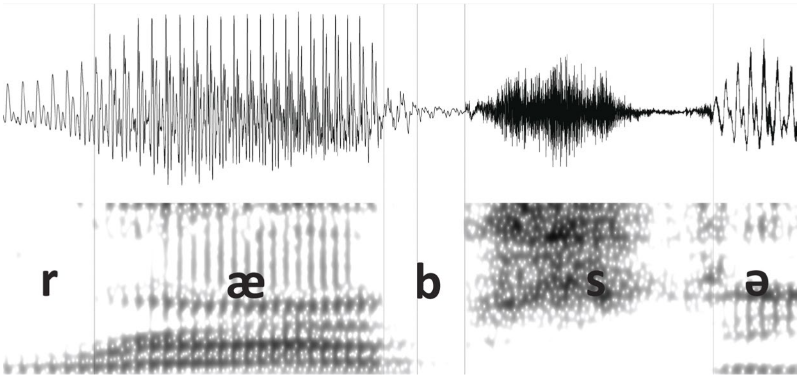

Figure 10. Segmentation example on rhapsody (rha_13_low.textgrid)

Figure 11 below shows a typical segmentation of *knapsack*. The preconsonantal vowel has a duration of 96 ms, followed by a 96 ms postvocalic stop. The duration of the intervocalic consonant cluster is 223 ms, whereas the combination of the vowel and consonant cluster is 319 ms, with a total duration of 641 ms for the entire occurrence. There was no voicing detected in the stop segment, thus marked "VLESS." The ratios calculated from these measurements include a vowel-to-consonant cluster duration ratio of 0.30 and an occurrence-to-vowel duration ratio of 0.15, with no voicing ratio calculated in the lack of a (mostly) regularly repeating pattern in the stop's acoustic waveform.

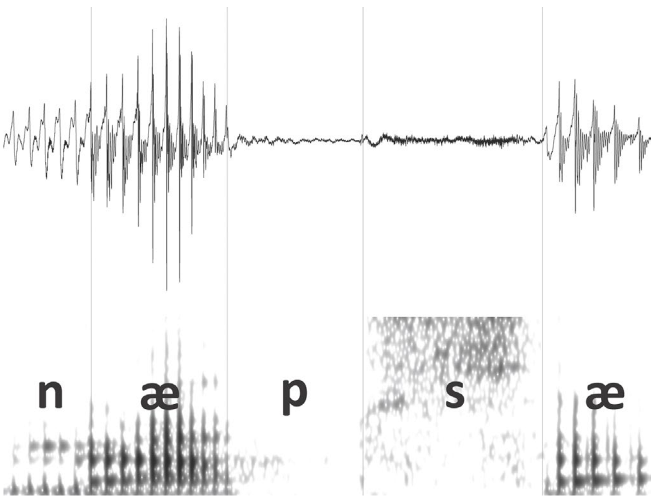

Figure 11. Segmentation example on knapsack (kna_20_low.textgrid)

4.2 Segment Duration

4.2.1 Duration of the Preconsonantal Vowel

Table 7 below summarizes the descriptive statistics for the preconsonantal vowel in the three test items, which is followed by the distribution and density of the data in Figure 12.

Item	n	Mean	Median	Min	Max	SD	SE
absolute	39	114	116	63	160	24.32	3.95
rhapsody	36	130	132.5	58	197	35.45	5.99
knapsack	38	118	115	71	163	20.29	3.33

Table 7. Duration of the preconsonantal vowel (in ms)

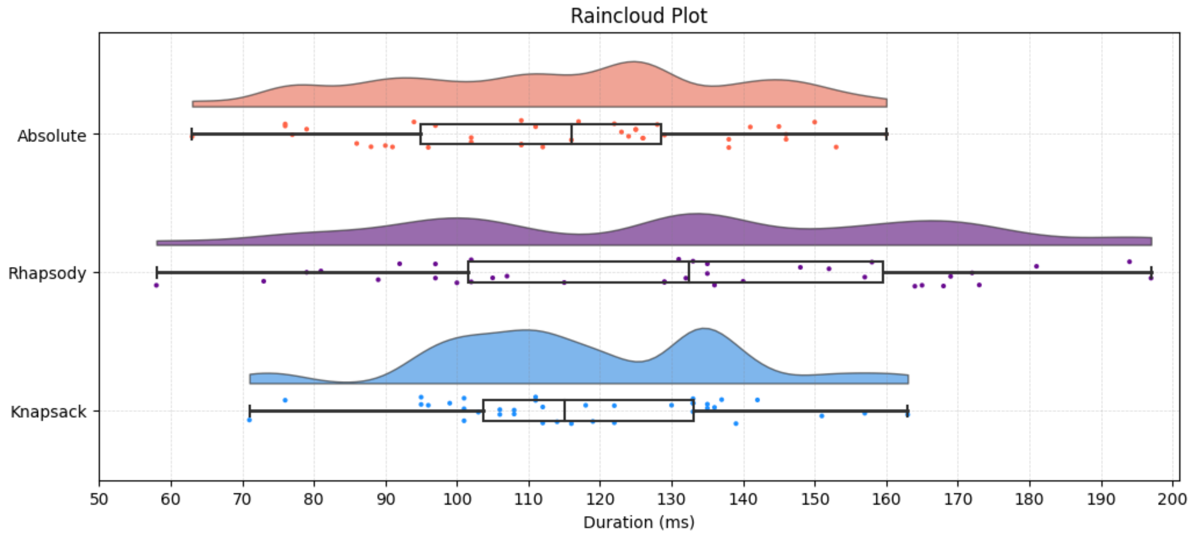

Figure 12. Distribution and density of the preconsonantal vowel (in ms)

The Kruskal-Wallis test yielded an H-statistic of 4.76 with a p-value of 0.0924, indicating that the differences between the groups are not statistically significant. The calculated η^2 was 0.0425, suggesting that about 4% of the variance in the data can be attributed to differences between the groups. The Mann-Whitney U test however revealed a significant difference between *absolute* vs. *rhapsody* ($p = 0.0313$) with a small to moderate effect size ($r = 0.3551$), indicating that the vowel in *rhapsody* was pronounced slightly slower. No statistically significant differences were found between *absolute* vs. *knapsack* ($p = 0.5786$, $r = 0.4629$), nor in *rhapsody* vs. *knapsack* ($p = 0.1398$, $r = 0.6001$), despite moderate to large effect sizes.

4.2.2 Duration of the Postvocalic Stop

Table 8 below summarizes the descriptive statistics for the postvocalic stop in the three test items, which is followed by the distribution and density of the data in Figure 13.

Item	n	Mean	Median	Min	Max	SD	SE
absolute	39	60	54	28	114	21.30	3.46
rhapsody	36	57	56.5	27	111	19.21	3.25
knapsack	38	69	70	35	102	17.43	2.86

Table 8. Duration of the postvocalic stop (in ms)

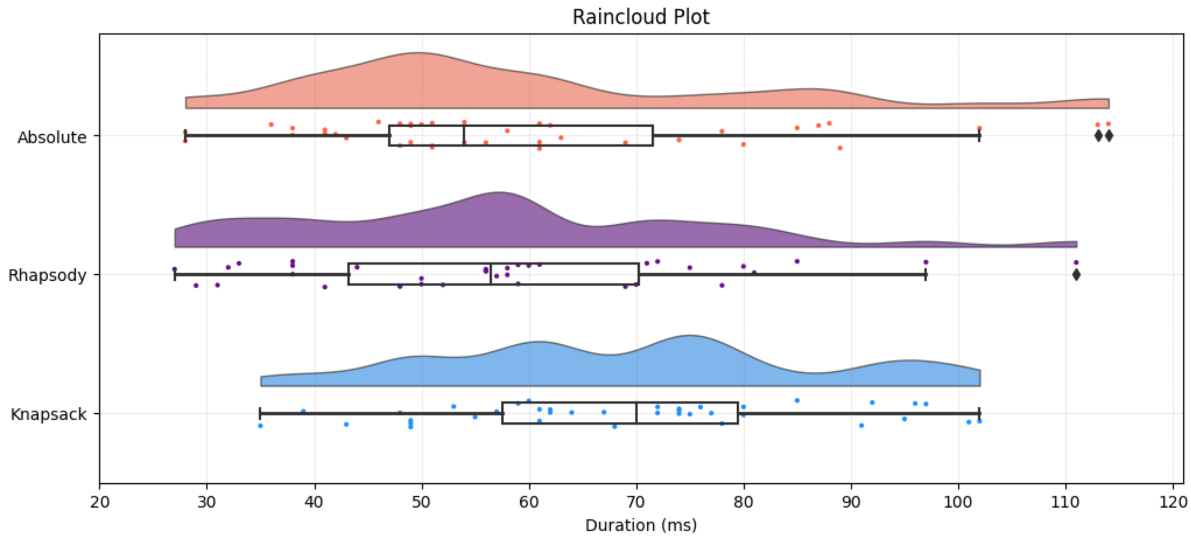

Figure 13. Distribution and density of the postvocalic stop (in ms)

The Kruskal-Wallis test yielded an H-statistic of 8.94 with a p-value of 0.0115, indicating that there are statistically significant differences between at least some of the groups. The calculated $\eta^2 = 0.0798$, that is, about 8% of the variance in the data can be attributed to differences between the groups. The Mann-Whitney U test revealed no statistically significant difference between *absolute* vs. *rhapsody* ($p = 0.7941$), although the effect size ($r = 0.518$) indicated a large effect. A statistically significant difference was however found between *absolute* vs. *knapsack* ($p = 0.0168$, $r = 0.341$) and similarly in *rhapsody* and *knapsack* ($p = 0.0061$, $r = 0.314$) both indicating a medium effect, and finally highlighting that the stop consonant of *knapsack* was pronounced with a longer stop-closure duration.

4.2.3 Duration of the Intervocalic Consonant Cluster

Table 9 below summarizes the descriptive statistics for the intervocalic consonant cluster in the three test items, which is followed by the distribution and density of the data in Figure 14.

Item	n	Mean	Median	Min	Max	SD	SE
absolute	39	176	171	108	284	37.15	6.03
rhapsody	36	155	158	108	227	25.67	4.34
knapsack	38	183	185.5	123	251	28.07	4.61

Table 9. Duration of the intervocalic consonant cluster (in ms)

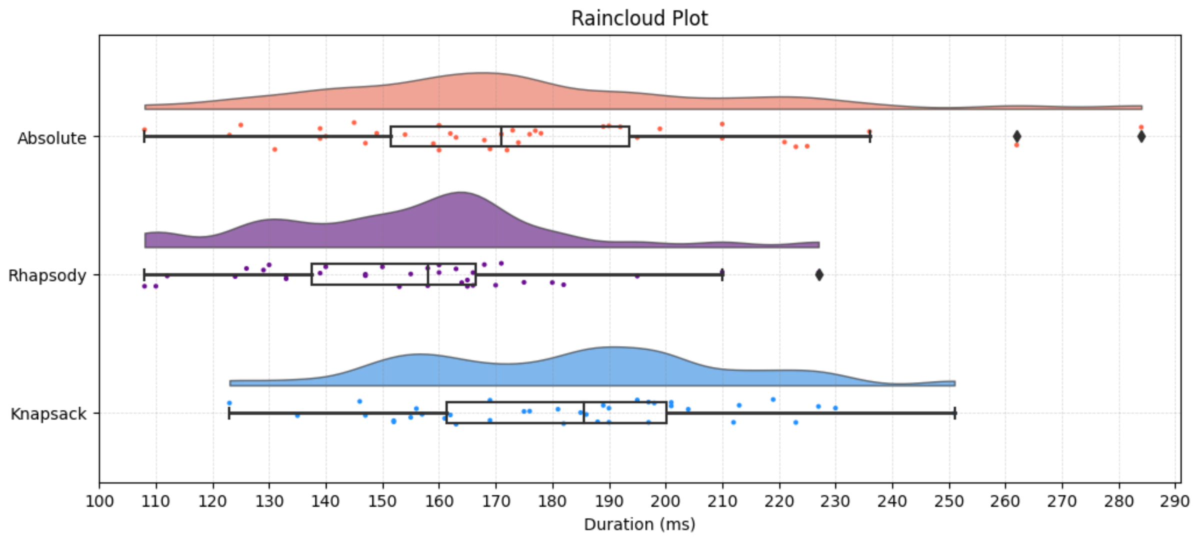

Figure 14. Distribution and density of the intervocalic consonant cluster (in ms)

The Kruskal-Wallis H-test yielded an H-statistic of 15.31 with a p-value of 0.00047, indicating statistically significant differences between at least some of the groups. The calculated $\eta^2 = 0.137$, that is, about 14% of the variance in the data can be attributed to differences between the groups. The Mann-Whitney U test revealed a significant difference between *absolute* vs. *rhapsody* ($p = 0.0134$) with a substantial effect size ($r = 0.666$). No statistically significant difference was found between *absolute* vs. *knapsack* ($p = 0.1836$), although the effect size ($r = 0.412$) indicated a medium effect. A statistically significant difference was found in *rhapsody* vs. *knapsack* ($p = 0.0001$) with a small to medium effect size ($r = 0.234$), which overall means that *rhapsody* had the shortest consonant cluster duration.

4.2.4 Duration of the Vowel and Consonant Cluster

Table 10 below summarizes the descriptive statistics for the vowel+consonant cluster in the three test items, which is followed by the distribution and density of the data in Figure 15.

Item	n	Mean	Median	Min	Max	SD	SE
absolute	39	290	294	184	393	44.23	7.17
rhapsody	36	286	274.5	202	400	50.53	8.54
knapsack	38	301	301.5	218	366	34.37	5.65

Table 10. Duration of the vowel+consonant cluster (in ms)

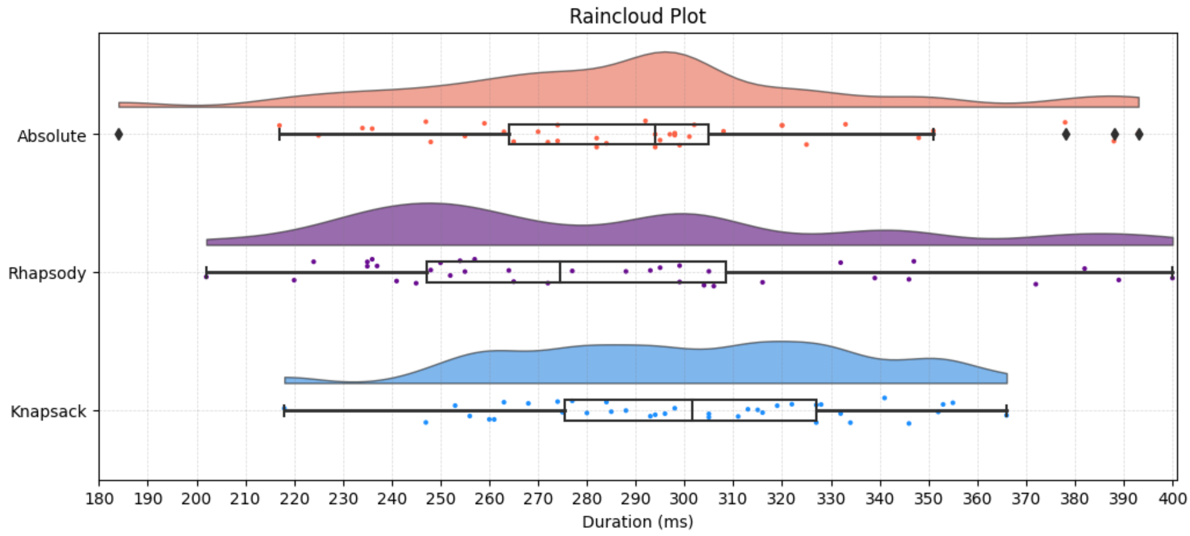

Figure 15. Distribution and density of the vowel+consonant cluster (in ms)

The Kruskal-Wallis test yielded an H-statistic of 4.01 with a p-value of 0.1346, indicating that the differences between the groups are not statistically significant. The calculated $\eta^2 = 0.0358$, that is, about 4% of the variance in the data can be attributed to differences between the groups. The Mann-Whitney U test revealed no statistically significant difference between *absolute* vs. *rhapsody* (p = 0.4939), although the effect size (r = 0.5463) indicated a large effect. No statistically significant difference was found between *absolute* vs. *knapsack* (p = 0.1673, r = 0.5463), and again between *rhapsody* vs. *knapsack* (p = 0.05767, r = 0.3713), despite the groups indicating a substantial and moderate effect sizes, respectively. The pairwise comparisons displayed a borderline non-significant difference among the three item groups.

Table 11 below summarizes the descriptive statistics for the ratio of the vowel's duration to the consonant cluster's duration in the three test items, which is followed by the distribution and density of the data in Figure 16.

Item	n	Mean	Median	Min	Max	SD	SE
absolute	39	0.67	0.66	0.27	1.16	0.19	0.03
rhapsody	36	0.85	0.85	0.35	1.34	0.23	0.04
knapsack	38	0.39	0.40	0.26	0.47	0.05	0.01

Table 11. Ratio of the vowel's duration to the consonant cluster's duration

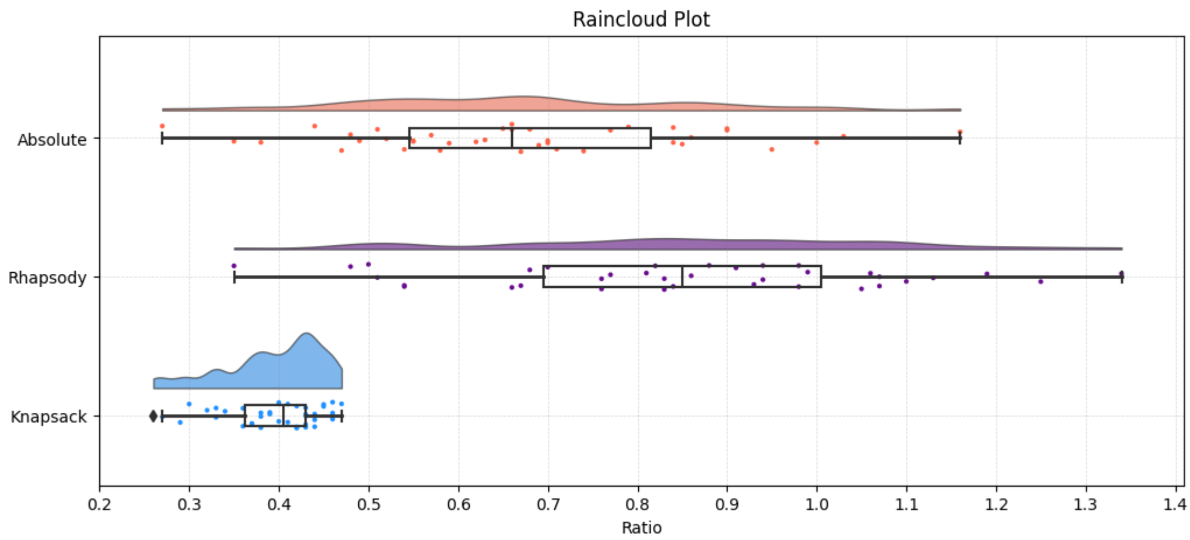

Figure 16. Ratio of the vowel's duration to the consonant cluster's duration

The Kruskal-Wallis test yielded an H-statistic of 67.86 with a p-value of 1.84×10^{-15}, indicating that there are statistically highly significant differences between at least some of the groups. The calculated $\eta^2 = 0.606$, that is, about 61% of the variance in the data can be attributed to differences between the groups. The Mann-Whitney U test revealed a significant difference between *absolute* vs. *rhapsody* ($p = 0.00105$) with a small to moderate effect size ($r = 0.280$). A statistically highly significant difference with a very strong effect size was however found between *absolute* vs. *knapsack* ($p = 6.20 \times 10^{-11}$, $r = 0.933$) and similarly between *rhapsody* vs. *knapsack* ($p = 1.55 \times 10^{-12}$, $r = 0.978$). Therefore, the vowel's duration in *knapsack* shows a statistically highly significant difference compared to the other two items.

4.2.5 Total Duration of the Occurrence

Table 12 below summarizes the descriptive statistics for the total occurrence duration in the three test items, which is followed by the distribution and density of the data in Figure 17.

Item	n	Mean	Median	Min	Max	SD	SE
absolute	39	502	491	281	711	88.66	14.38
rhapsody	36	559	562	352	785	115.06	19.45
knapsack	38	627	624	474	791	82.47	13.56

Table 12. Total duration of the occurrences (in ms)

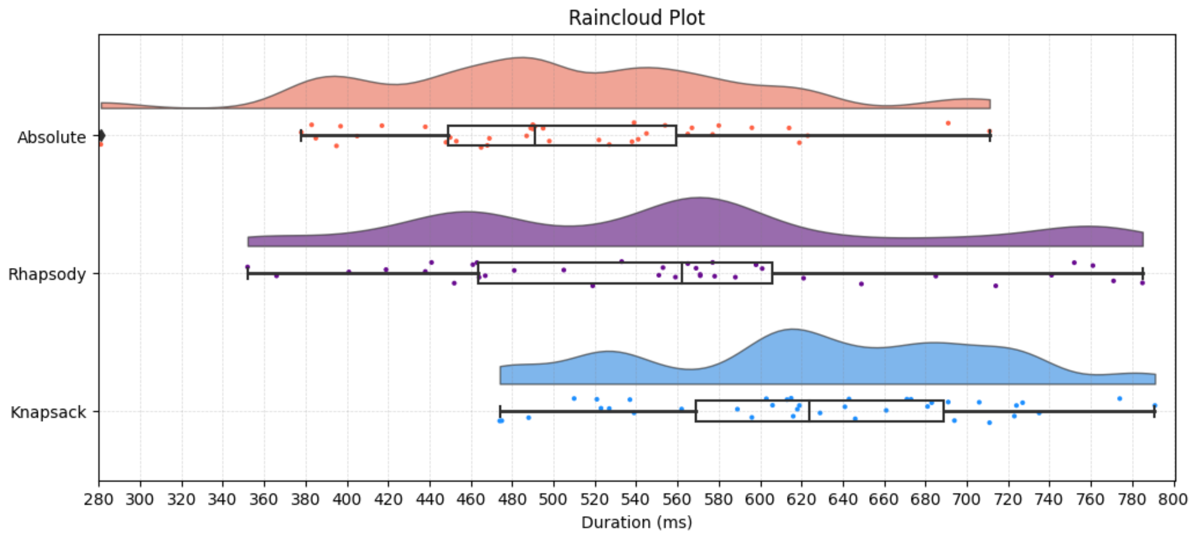

Figure 17. Distribution and density of the occurrences (in ms)

The Kruskal-Wallis test yielded an H-statistic of 25.98 with a p-value of 2.28e-06, indicating that there are statistically significant differences between at least some of the groups. The calculated $\eta^2 = 0.232$, that is about 23% of the variance in the data can be attributed to differences between the groups. The Mann-Whitney U test revealed a significant difference between *absolute* vs. *rhapsody* (p = 0.0439) with a medium effect size (r = 0.3643). A statistically highly significant difference was however found between *absolute* vs. *knapsack* (p = 3.22e-07) with an effect size (r = 0.1613) indicating a relatively small effect, and similarly between *rhapsody* vs. *knapsack* (p = 0.0044) with an effect size (r = 0.3070) displaying a moderate effect. With this in mind, the occurrences were pronounced with varying durations.

Table 13 below summarizes the descriptive statistics for the ratio of the vowel's duration to the total occurrence's duration in the three test items, which is followed by the distribution and density of the data in Figure 18.

Item	n	Mean	Median	Min	Max	SD	SE
absolute	39	0.23	0.23	0.12	0.33	0.05	0.01
rhapsody	36	0.24	0.24	0.13	0.40	0.06	0.01
knapsack	38	0.19	0.19	0.12	0.24	0.03	0.01

Table 13. Ratio of the vowel's duration to the total occurrence's duration

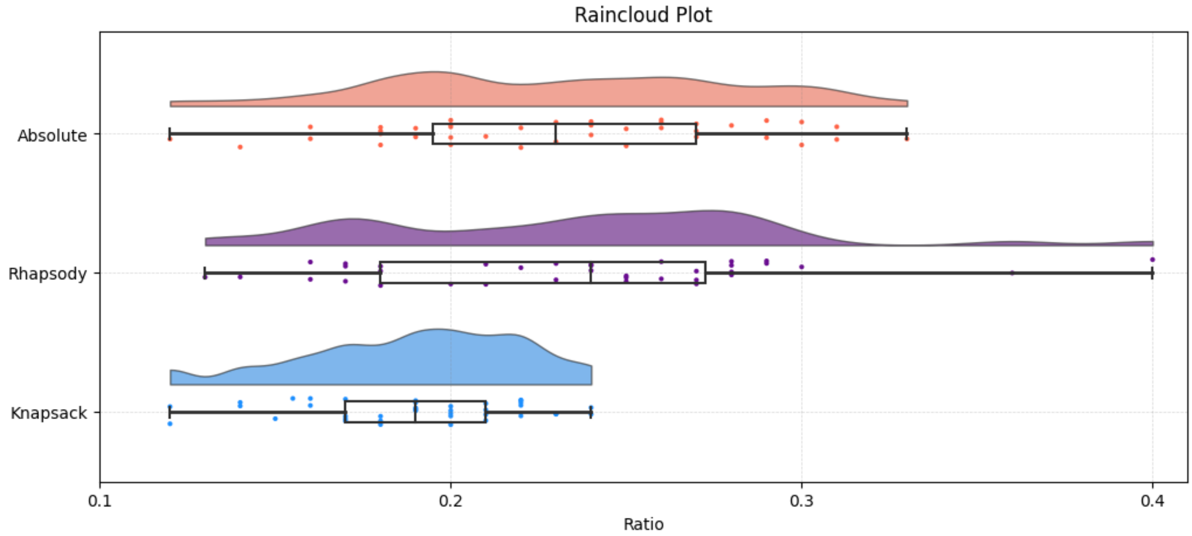

Figure 18. Ratio of the vowel's duration to the total occurrence's duration

The Kruskal-Wallis test yielded an H-statistic of 18.09 with a p-value of 0.00012, indicating that there are statistically significant differences between at least some of the groups. The calculated $\eta^2 = 0.162$, that is, about 16% of the variance in the data can be attributed to differences between the groups. The Mann-Whitney U test revealed no significant difference between *absolute* vs. *rhapsody* ($p = 0.85$), despite the effect size ($r = 0.487$) suggesting a medium to large difference. A statistically highly significant difference with an effect size indicating a strong effect was found between *absolute* vs. *knapsack* ($p = 0.00026$, $r = 0.741$), and similarly between *rhapsody* vs. *knapsack* ($p = 0.00023$, $r = 0.749$), indicating that the vowel length of *knapsack* was statistically shorter than that of *absolute* and *rhapsody*.

4.3 Stop Voicing

Table 14 below summarizes the descriptive statistics of stop voicing in the three test items, which displays that a substantial proportion of *absolute* (92%) and *rhapsody* (89%) contained some voicing into the stop consonant, whereas in *knapsack* almost $^2/_3$ of the tokens (61%) were voiceless, indicating that only a moderate proportion contained some voicing.

Item	n	Criteria	Proportion
absolute	**36**	VD	0.92
absolute	3	VLESS	0.08
rhapsody	**32**	VD	0.89
rhapsody	4	VLESS	0.11
knapsack	**15**	VD	0.39
knapsack	23	VLESS	0.61

Table 14. Voicing in the stop consonant

From this point on, only those stops were considered that had some voicing in their stop closure.

Item	n	Mean	Median	Min	Max	SD	SE
absolute	**36**	35	33	9	72	15.56	2.63
rhapsody	**32**	28	22	9	111	18.95	3.40
knapsack	**15**	24	18	11	69	16.92	4.52

Table 15. Duration of voicing in the stop consonant (in ms)

Item	n	Mean	Median	Min	Max	SD	SE
absolute	**36**	0.62	0.60	0.18	1.00	0.24	0.04
rhapsody	**32**	0.49	0.45	0.16	1.00	0.23	0.04
knapsack	**15**	0.33	0.26	0.12	0.74	0.19	0.05

Table 16. Voicing ratio in the stop consonant

Table 15 above summarizes the descriptive statistics for voicing duration, and Table 16 displays the voicing ratio in the stop consonants' closure interval (or in some cases even farther), which is followed by the distribution and density of the data in Figure 18 and 19.

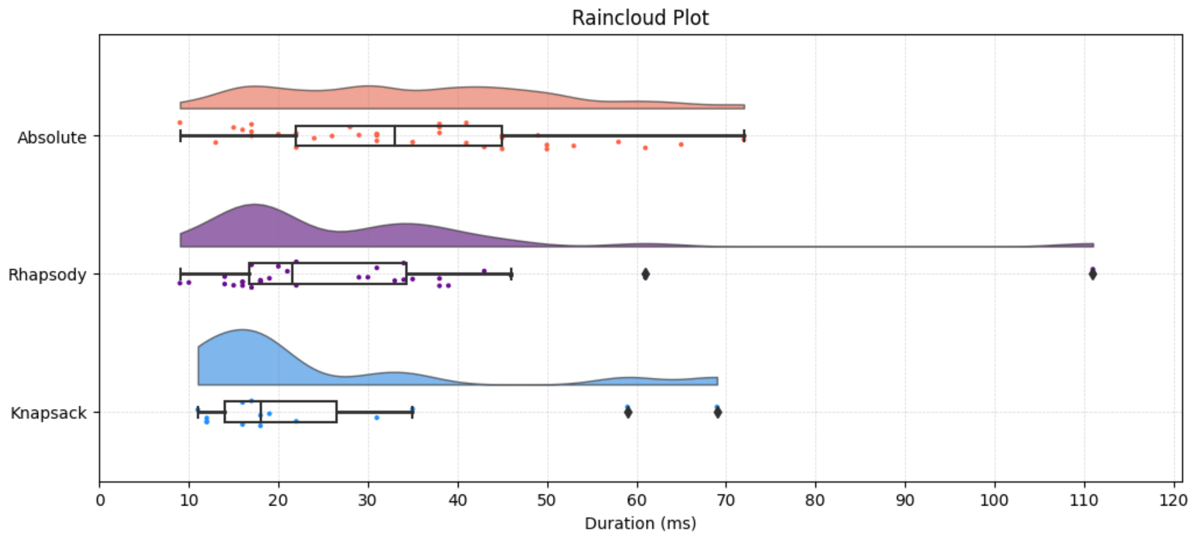

Figure 19. Duration of voicing in the stop consonant (in ms)

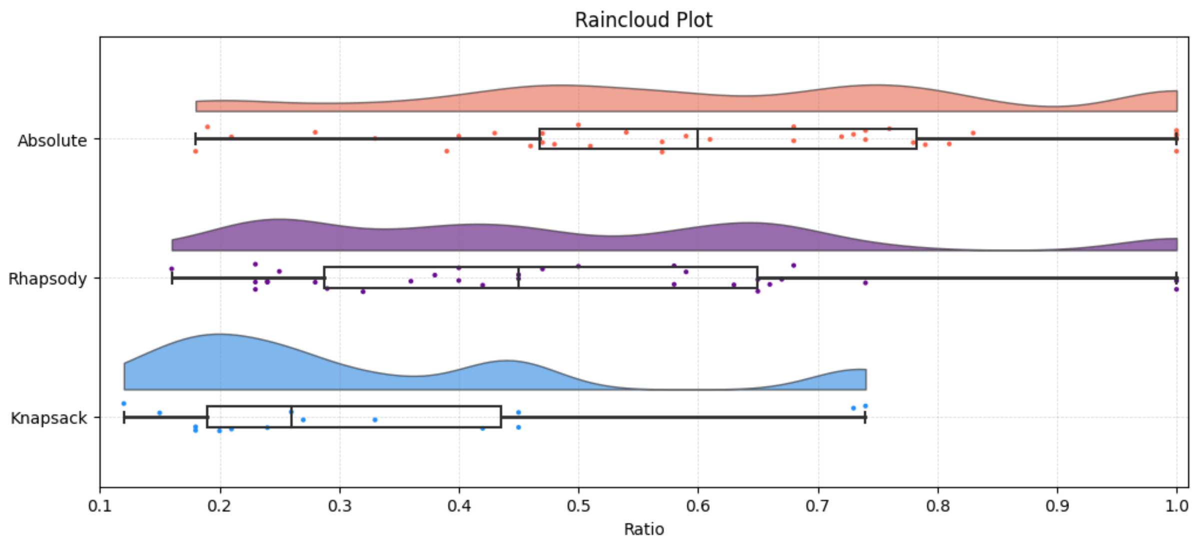

Figure 20. Voicing ratio in the stop consonant

For Figure 19, the Kruskal-Wallis test yielded an H-statistic of 8.80 with a p-value of 0.012, indicating that there are statistically significant differences among the groups. The calculated $\eta^2 = 0.1073$, that is, about 11% of the variance in the data can be attributed to differences between the groups. The Mann-Whitney U test revealed a significant difference between *absolute* vs. *rhapsody* ($p = 0.024$) with a strong effect size ($r = 0.660$). A statistically significant difference was found between *absolute* vs. *knapsack* ($p = 0.013$) with an effect size ($r = 0.724$) indicating a substantial effect. No statistically significant differences were however found between *rhapsody* vs. *knapsack* ($p = 0.253$), despite a large effect size ($r = 0.605$).

For Figure 20, the Kruskal-Wallis test yielded an H-statistic of 16.31 with a p-value of 0.00029, indicating that there are statistically significant differences among the groups. The calculated $\eta^2 = 0.1989$, that is, about 20% of the variance in the data can be attributed to differences between the groups. The Mann-Whitney U test revealed a significant difference between *absolute* vs. *rhapsody* ($p = 0.021$) with a strong effect size ($r = 0.664$). Though, a statistically highly significant difference was found between *absolute* vs. *knapsack* ($p = 0.00018$) with an effect size ($r = 0.836$) indicating a very strong effect, and similarly between *rhapsody* vs. *knapsack* ($p = 0.018$) with an effect size ($r = 0.717$) displaying a very large effect. As follows, there is significantly higher degree of voicing in *absolute* and *rhapsody* than in *knapsack*.

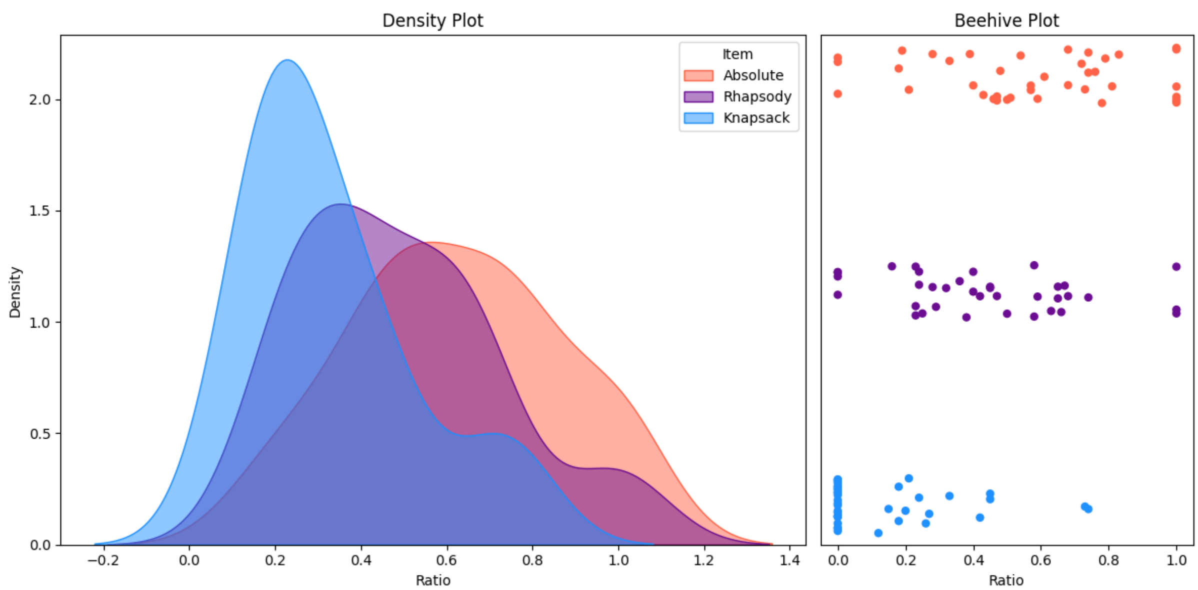

Figure 21. Overlapping density plot and beehive plot for voicing ratio

On the left-hand side in Figure 21 above, the voicing ratio in the stop consonant is summarized with overlapping density curves, highlighting that *knapsack* had the least amount of stop voicing. On the right hand side in Figure 20 above, this beehive plot serves to represent both the 'voiced' and 'voiceless' stop consonants to show the overall distribution and ratio of voicing in the investigated stop consonants, similarly displaying that *absolute* and *rhapsody* had considerably more and longer stop voicing than *knapsack*.

5 DISCUSSION

5.1 Interpretation of Segment Durations

The results from the experiment and the statistical analysis across the three test items indicate significant differences only in the duration of the postvocalic stop. The stop consonant in *knapsack* was on average 9 ms longer than in *absolute* and 12 ms longer than in *rhapsody*, which points to the phonetic fact argued by Bickford (1985: 202) that fortes usually exhibit longer consonant durations (and are also likely to be louder) than their lenis counterparts.

The vowel+consonant cluster and total occurrence durations indicate borderline non-significant differences across the items, likely due to the generated YouGlish videos which randomly contained words with differing pronunciation rates and segmental durations. These complications (individual differences) may be responsible for by-chance longer vowel and shorter intervocalic consonant cluster realizations in the *rhapsody* data. The complication is addressed by the process of calculating ratios, which have been found to remain "relatively invariant" across variations in speaking rate, for example (G. Kiss 2013: 11). The ratio of the preconsonantal vowel's duration to consonant cluster's duration indicates that on average *knapsack* (0.39) has a significantly smaller ratio compared to *absolute* (0.67) and *rhapsody* (0.85). This finding is further supported by the ratio of the vowel's duration to the occurrence's duration, respectively indicating ~21% and ~26% longer vowels in *absolute* and *rhapsody*, which underscore the well-known phenomenon of prefortis clipping, evident in *knapsack*, according to which vowels are usually realized shorter/clipped before fortes compared to lenes.

Experimental evidence thus confirms that the underlying identity of *knapsack*'s stop consonant appears to be a fortis /p/, as is standardly assumed. In contrast, the stop in *rhapsody* (despite conventional analysis), just like in *absolute*, lacks indispensable cues of a clipping effect and consistently displays shorter duration, encoding its lenis nature. Thus, it is reasonable to assume that both *absolute* and *rhapsody* can be analyzed with a LF, that is, /bs/ cluster.

5.2 Interpretation of Stop Voicing

The results from the stop-voicing statistics further support our hypothesis that both *absolute* and *rhapsody* contain an underlying LF, that is, /bs/ cluster, while *knapsack* contains a morphologically complex FF, that is, /ps/ cluster. The stop consonant of *absolute* and *rhapsody*, in 92% and 89% of the cases respectively, revealed some simple, low frequency, quasiperiodic vocal fold vibrations in the waveform. Contrarily, *knapsack* contained no such vibrations about two out of three times or 61% of the cases, rendering it usually fully voiceless, but even so, as many as 15 tokens in the *knapsack* data did exhibit some voicing extending into the stop.

The voiced tokens of *knapsack* displayed a relatively compact distribution with a mean length of 24 ms. This suggests that the voicing observed is "usually indicative of coarticulatory voicing, i.e., voicing is just a by-product residue of the preceding vowel's final few vocal fold pulsations," which means that voice passes through the walls of the vocal tracts and continues to shape the waveform (G. Kiss & Szigetvári 2020: 151; Mannell 2024: 4).

The voiced tokens of *absolute* and *rhapsody* had longer average voicing durations, with 35 ms and 28 ms respectively, when compared to the shorter 24-ms duration in *knapsack*. Although the statistical analysis did not reveal significant differences in terms of voicing duration between *rhapsody* and *knapsack*, the mean duration of *knapsack* may be right-skewed, likely due to two outlier pronunciations with 69 ms and 59 ms in the extremes, making its duration potentially similar to *rhapsody*'s mean score. Similarly, these extremes may contribute to the overall 'higher' mean voicing ratio of 0.33 given their ratios of 0.73 and 0.74. These cases may indicate some unexpected planned voicing articulations, that is, some deviant realizations, suggesting that the stop consonant may even resemble a lenis /b/ in these instances.

Nonetheless, the calculated voicing ratios revealed that 4 cases in *absolute* and 3 cases in *rhapsody* display 1:1 voicing ratios (voicing duration relative to the stop's duration), which indicate that these words most likely prompted some planned voicing from the speakers.

Moreover, the mean voicing ratios of *absolute* (0.62) and *rhapsody* (0.49) show a passive implementation of voicing (cf. Docherty 1992: 120), despite the lack of an intersonorant position, which also confirms Iverson and Salmons' (2003: 53) claim that (even final) "lenis obstruents may acquire some measure of voicing from a preceding voiced segment." Thus, lenes may be partially voiced in languages without active voicing. Although some voicing is present in the *knapsack* data, the stop predominantly has a voiceless articulation, which is supported, if Bárkányi and Mády's (2012) findings are applicable to English, by the fact that only 6 items exceeded the identified ~30% threshold for an obstruent to be classified as voiced.

Besides this, the statistical analysis indicated that there is a highly significant difference in terms of voicing ratio in *absolute* vs. *knapsack* and in *rhapsody* vs. *knapsack*. Though, a statistically significant difference was also found between *absolute* vs. *rhapsody*, possibly due to the overall higher voicing duration in the stop consonant in the *absolute* data. The reasons for this may be varied. The word *rhapsody* originates from Ancient Greek, a language whose spelling system marks laryngeal assimilation at morpheme boundaries (root consonant assimilation in this case), rather than closely reflecting morphemes. In *ῥαψῳδία* ('rhapsody'), the root *ῥαπ-* (from *ῥάπτω* 'to sew') is combined with *ᾠδή* ('song'). This cluster in *ῥαψῳδία* is represented by the letter *ψ* (psi), the combination of /p/ and /s/, which indicates that the aspirate /pʰ/ and the etymological root /b/ are neutralized before /s/, thus English ultimately borrows this word with its Latinized spelling <ps>. Contrary, the word *absolute* originates from Latin, a language whose spelling system kept the integrity of morphemes in some prefixes, rather than being closely phonetic. This word derives from *absolutus* (the past participle of *absolvere* 'to set free' or 'to complete,' combining the prefix *ab-* 'from' and the root *solvere* 'to loosen, release'). English ultimately inherits this word with its original Latin spelling. And so, it seems that our transcription conventions reflect etymology in both *absolute* (/bs/) and *rhapsody* (/ps/), which may be the result of a mismatch between spelling and phonology given our data supporting their identical underlying reality as /bs/. Regardless, speakers may feel the need to make a difference based on spelling, <b> vs. <p>, leading to a scalar distribution between the stop of *absolute* vs. *rhapsody*. Accordingly, we observe a longer voicing duration and larger voicing ratio in our *absolute* data. Seeing a <b> in *absolute* may trigger some voicing, while a <p> in *rhapsody* may not bring about a large effect, where spelling (alone) cannot mislead speakers into pronouncing a moderately voiced stop. On this basis, other phonetic cues may be more liable for the contrast, which is the reason why we found a shorter voicing duration and a smaller voicing ratio in our *rhapsody* data.

Nevertheless, if spelling were truly an influencing factor, then we would consistently find the stop pronounced with a completely voiceless interval (without any vocal fold vibrations). Our data only shows 4 tokens matching this criterion for *rhapsody*. This means that regardless of spelling, speakers associate this <p> with a lenis /b/ and produce it with some voicing during the occlusion. Given this, a lenis /b/ in *absolute* and *rhapsody* may show some planned voicing [b̥] but may also have the same phonetic imprint as an unaspirated fortis [p].[34]

As observed, *absolute* and *rhapsody* contain significantly higher proportions of stop voicing than *knapsack*. Both *absolute* and *rhapsody* display longer stop voicing, which is also considerably reflected in the calculated voicing ratios. Voicing values appear to serve as an additional appreciable phonetic cue that help distinguish if a stop is fortis or lenis. Our findings reflect this phenomenon and are in line with G. Kiss and Szigetvári (2020: 151) who discuss that "fortis obstruents are only voiced phonetically due to coarticulation [but] lenis obstruents may be phonetically voiced through allowing voicing to continue longer from the preceding vowel." Thus, *absolute* and *rhapsody* seem to share an identical LF, that is, /bs/ cluster, while *knapsack* has a morphologically complex FF, that is, /ps/ cluster based on our voicing facts.

[34] The reason why German lenis /b, d, g/ were borrowed as a voiceless /p, t, k/ into Hungarian, cf. ***b**inder*–***p**inét*, *Leopol**d**–Lipó**t***, ***g**aukler*–***k**ókler*. Interestingly, French ***pité*** was borrowed as a fortis /p/ in ***p**ity* (not ***b**ity*), which likely originated with the French aristocracy, and by the time it reached the public its spelling was already set.

5.3 Overall Implication of the Experiment

Experimental evidence supports the hypothesis, which posits that monomorphemic FF clusters are ruled out in English. Our findings demonstrate that the postvocalic stop in *knapsack* is a fortis /p/, while *absolute* and *rhapsody* share the same underlying lenis /b/. The observed segment and voicing durations, along with the calculated ratios all point toward this conclusion.

The results indicate that the stop consonants in *absolute* and *rhapsody* were articulated with a significantly shorter interval, lacked clipping effect (evidenced by consistently longer vowel durations when compared to *knapsack*, though the /æ/ in both *absolute* and *rhapsody* may still be shorter due to a following fortis /s/, cf. *tabs on*, where /æ/ may be considerably longer), and showed significantly higher proportions and longer durations of stop voicing, which was beyond coarticulatory voicing. These findings compel us to challenge standard transcription conventions (and possibly the prevailing view among many analysist who subscribe to traditional analysis) in proposing that both *absolute* and *rhapsody* contain an underlying LF, that is, /bs/ cluster. On the contrary, there is sufficient proof to support the standard view that *knapsack* contains a fortis stop as part of a morphologically complex F#F /p#s/ cluster. This is evidenced by its shorter vowel duration (indicative of prominent prefortis clipping) and the prevalence of a larger proportion of fully voiceless and longer stop intervals.

Therefore, the observed differences underline the role of phonetic cues in distinguishing lenis from fortis stops (or vice versa), especially in cases where orthography may (or may not) influence transcription conventions. As discussed, (etymological) spelling, which is reflected in transcriptions, cf. *absolute* (standardly with a LF /bs/ cluster) and *rhapsody* (standardly with a FF /ps/), possibly slightly mislead speakers into articulating the stop of *absolute* with significant planned voicing. Nevertheless, the stop of *rhapsody* still showed a considerable proportion and longer duration of voicing when adjusted to the *knapsack* data (which may be skewed due to the two discussed outlier pronunciations), despite an 'apparent' fortis <p> in *rhapsody*'s spelling, which highlights the fact that voicing distinctions may not be encoded in alphabetic scripts and are alone not necessarily a deciding factor to which series an obstruent in obstruent clusters belongs in a fortis-lenis language such as English.

Lastly, some caveats are in order. These findings may have broader implications for our understanding of the interaction across spelling and pronunciation in aspirating languages, suggesting that speakers may subconsciously align their articulation with orthographic cues, even when these cues do not perfectly match the underlying phonological representation, as per the voicing facts in the *absolute* data. Nonetheless, this research may have several limitations. The use of YouGlish for data collection introduces variability due to differing speaking rates and segmental durations, which as noted above may have impacted some of the results. The relatively small number of tokens analyzed for each word may limit generalizations, and not all measurable phonetic differences may be perceptually relevant.

Given the findings and the potential limitations of the experiment, some areas warrant further investigation as part of a comprehensive research endeavor or a PhD dissertation. As the underlying identity of obstruents in obstruent clusters has not been fully addressed in the literature, future research is much needed in a controlled environment to explore with a large-scale data set and a wider range of words the underlying identity of these clusters, for example found in *Ibsen–Ypsilanti–lipstick*, *mobster–synopsis–chopstick*, or *webster–Pepsi–prepschool*, while keeping in mind additional variables such as suprasegmentals to further validate the conclusions drawn here. Analyzing similar patterns and cluster behavior in various Germanic languages may further help generalize the findings beyond English and confirm the robustness of the hypothesis, which potentially excludes FF clusters from English phonology.

All things considered, our experimental evidence leads us to propose that both *absolute* and *rhapsody* feature the same underlying LF, that is, /bs/ cluster, thus /ǽ**bs**əlut/ and /rǽ**bs**ədɪ/.

6 CONCLUSION

This thesis has provided a theoretical background and empirical experimentation into voiceless obstruent clusters, focusing on the *fortis+fortis (*FF) phonotactic constraint, which posits that due to articulatory constraints adjacent fortis obstruents are disallowed in monomorphemic words in English. This thesis aimed to test the validity and the rigorousness of the *FF hypothesis and sought to potentially offer a reevaluation of how these clusters may be better represented in the phonology of Standard English.

The investigation of the orthographic <abs> vs. <aps> sequences in *absolute*, *rhapsody*, and *knapsack* validate the *FF hypothesis by reinforcing our as-if assumption. Accordingly, *rhapsody* may in fact be better represented as a LF, that is, /bs/ cluster (despite traditional analyses claiming it to be a FF, that is, /ps/ cluster) comparable to the standardly considered LF, that is, /bs/ cluster in *absolute*. This is evidenced by segment durations and ratios, as well as voicing proportions and ratios observed in our experimental data, thus *rhapsody* /rǽ**bs**ədɪ/ and *absolute* /ǽ**bs**əlut/ both contain a postvocalic lenis /b/, which distinguishes them from the fortis /p/ of the morphologically complex FF cluster in our *knapsack* /nǽ**ps**æk/ data.

It is our suggestion that standard transcription practices are fraught with difficulties and cannot cope with the underlying reality of *rhapsody*, which appears to contain a lenis /b/, because they are heavily influenced by orthographic conventions. This thesis suggests that a more phonologically-minded approach to transcriptions (especially in light of the justifiable *FF phonotactic constraint) may prevent misanalysis and lead to a more accurate underlying representation of obstruents in obstruent clusters in English.

These findings possibly open up the way for further research and empirical analysis with a larger dataset, investigating voiceless obstruent clusters not only in Standard English, but also in ‘peculiar’ varieties of English and various other aspirating Germanic languages, which could test the robustness of *FF and further explore its applicability cross-linguistically.

Finally, this thesis has aimed to contribute to the ongoing discourse in phonological theory by providing experimental evidence that supports the *FF phonotactic constraint in *rhapsody* and by advocating for a phonologically-minded approach for transcribing voiceless obstruent clusters, which underlines the importance of aligning phonological theory with experimental evidence in differentiating fortis/lenis segments in obstruent clusters in English.

7 REFERENCES

Abramson, Arthur S. & Whalen D. H. 2017. Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics* 63(1). 75–86. https://doi.org/10.1016/j.wocn.2017.05.002.

Allen, Micah & Davide Poggiali, Kirstie Whitaker, Tom Rhys Marshall, Jordy van Langen, Rogier A. Kievit. 2019. Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Res.* 4(63). https://10.12688/wellcomeopenres.15191.2.

Aoun, Youssef. 1979. Is the syllable or the supersyllable a constituent? *MIT Working Papers in Linguistics* 1(1). 140–148.

Avery, Peter & William J. Idsardi. 2001. Laryngeal dimensions, completion and enhancement. In Tracy Alan Hall (edn.), *Distinctive feature theory*. Berlin: Mouton de Gruyter. 41–70.

Bárkányi, Zsuzsanna & Katalin Mády. 2012. The perception of voicing in fricatives. Paper presented at the *9th Old World Phonology Conference*, 01.18.2012, Berlin, Germany.

Backley, Phillip. 2011. *Introduction to Element Theory*. Edinburgh University Press.

Beckman, Jill & Michael Jessen, Catherine Ringen. 2013. Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics* 49(2). 259–284. https://doi.org/10.1017/S0022226712000424.

Beňuš, Štefan. 2021. *Investigating Spoken English*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-54349-5.

Bickford, J. Albert. 1985. Fortis/lenis consonants in Guichicovi Mixe: A preliminary acoustic study. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session*. 29(1). 195–207. https://doi.org/10.31356/silwp.vol29.01.

Boersma, Paul & David Weenink. 2023. *Praat: Doing phonetics by computer*. [Computer program.] Version 6.3.18. http://www.praat.org/.

Carnegie Mellon University. 2014–. *The CMU Pronouncing Dictionary*. http://www.speech.cs.cmu.edu/cgi-bin/cmudict?in=C+M+U+Dictionary#about.

Catford, John Cunnison. 1977. *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.

Cho, Young-Mee Y. 1990. A typology of voicing assimilation. *West Coast Conference on Formal Linguistics* 9(1). 141–155.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. *Papers in Laboratory Phonology I: Between the grammar and the physics of speech.* 283–333.

Cruttenden, Alan. 2014. *Gimson's pronunciation of English*, 8th edition. London & New York, NY: Routledge. https://doi.org/10.4324/9780203784969.

Cyran, Eugeniusz. 2014. *Between phonology and phonetics: Polish voicing*. Boston & Berlin: De Gruyter Mouton.

Davidsen-Nielsen, Niels. 1969. English stops after initial /s/. *English Studies* 4(1). 321–339. https://doi.org/10.1080/00138386908597340.

Davidson, Lisa. 2016. Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics* 54. 35–50. https://doi.org/10.1016/j.wocn.2015.09.003.

Docherty, Gerard J. 1992. *The timing of voicing in British English obstruents*. Berlin: Walter de Gruyter.

G. Kiss, Zoltán. 2013. Measuring acoustic correlates of voicing in stops and fricatives. In Péter Szigetvári (edn.), *VLLXX: Papers Presented to László Varga on his 70th Birthday*. Budapest: Department of English Linguistics, Eötvös Loránd University & Tinta Publishing House. 289–311. http://seas.elte.hu/VLlxx/gkiss.html.

G. Kiss, Zoltán. 2017. Aspiration of stops after fricatives in English: Results from a pilot experiment. In Péter Szigetvári (edn.), *70 snippets to mark Ádám Nádasdy's 70th birthday.* Budapest: Department of English Linguistics, Eötvös Loránd University. http://seas.elte.hu/nadasdy70/gkiss.html.

G. Kiss, Zoltán & Péter Szigetvári. 2020. Telling fortis and lenis apart in English obstruent clusters. *The Even Yearbook* 14(1). Budapest: Department of English Linguistics, Eötvös Loránd University 135–156. http://seas.elte.hu/w/!even/_media/20gs.pdf.

G. Kiss, Zoltán. 2021–. *Phonolec notes*. [Unpublished manuscript.] Budapest: Department of English Linguistics, Eötvös Loránd University. http://www.budling.hu/~cash/courses/phonolec/phonolec-notes/index.html.

Goldstein, Louis. 2022a. *Prosody: Temporal structure*. [PowerPoint.] University of South Carolina. https://sail.usc.edu/~lgoldste/General_Phonetics/Slides/Prosodic%20Phrasing%20and%20Rhythm.pdf.

Goldstein, Louis. 2022b. *Voicing and aerodynamics*. [PowerPoint.] University of South Carolina. https://sail.usc.edu/~lgoldste/General_Phonetics/Slides/Voicing%20and%20aerodynamics.pdf

Google. 2024. Google Colaboratory. https://colab.research.google.com/.

Hall, Tracy Alan. 2001. Introduction: Phonological representations and phonetic implementation of distinctive features. In Tray Alan Hall (edn.), *Distinctive feature theory*. Berlin, Boston: Mouton de Gruyter Mouton. 1–40.

Halle, Mohanan & Kenneth N. Stevens. 1971. A note on laryngeal features. *MIT Research Laboratory of Electronics Quarterly Progress Report* 101. 198–213.

Harris, John. 1994. *English sound structure.* Oxford & Cambridge, MA: Blackwell.

Hogg, Richard M. 1992. *A grammar of Old English. Vol. 1: Phonology*. Oxford: Blackwell.

Honeybone, Patrick. 2005. Diachronic evidence in segmental phonology: The case of obstruent laryngeal specifications. In Marc van Oostendorp & Jeroen van de Weijer (edn.), *The internal organization of phonological segments*. Berlin & New York: Mouton de Gruyter. 319–354. https://doi.org/10.1515/9783110890402.317.

International Phonetic Association (IPA). 2015. *Full IPA Chart*. Creative Commons Attribution-Sharealike 3.0 Unported License.

Iverson, Gregory K. & Joseph C. Salmons. 1995. Aspiration and laryngeal representation in Germanic. *Phonology* 12(3). 369–396. https://doi.org/10.1017/S0952675700002566.

Iverson, Gregory K. & Joseph C. Salmons. 2003. Laryngeal enhancement in early Germanic. *Phonology* 20(1). 43–74. https://doi.org/10.1017/S0952675703004469.

Jansen, Wouter. 2007. Phonological 'voicing,' phonetic voicing, and assimilation in English. *Language Sciences* 29(2–3). 270–293. https://doi.org/10.1016/j.langsci.2006.12.021.

Jessen, Michael. 1998. *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam and Philadelphia: John Benjamins Publishing Co.

Jessen, Michael. 1999. *Phonetics and phonology of tense and lax obstruents in German*. Amsterdam and Philadelphia: John Benjamins Publishing Co.

Jessen, Michael & Catherine Ringen. 2002. Laryngeal features in German. *Phonology* 19(2). 1–30. https://doi.org/10.1017/S0952675702004311.

Jones, Daniel. 1917. *English Pronouncing Dictionary*. Dent.

Jones, Daniel. 1957. *An outline of English phonetics*. 8th edition. Cambridge: W. Heffer & Sons Ltd.

Kahn, Daniel. 1976. *Syllable-based generalisations in English phonology*. [Doctoral dissertation.] MIT, Cambridge, MA. Published by Garland Press, New York.

Kaye, Jonathan. 1992. Do you believe in magic? The story of s+C sequences. *SOAS Working Papers in Linguistics & Phonetics* 2(1). 293–313.

Keating, Patricia A. 1984. Phonetic and phonological representation of stop consonant voicing. *Language* 60(2). 286–319. https://doi.org/10.2307/413642.

Kenyon, John Samuel & Thomas Albert Knott. 1953. *A pronouncing dictionary of American English*. Springfield, Mass: G. & C. Merriam Co.

Kerswill, Paul E. 1987. Levels of linguistic variation in Durham. *Journal of Linguistics* 23(1). 25–49. Cambridge University Press. https://doi.org/10.1017/S0022226700011026.

Kim, Heejin & Jennifer Cole. 2005. The stress foot as a unit of planned timing: Evidence from shortening in the prosodic phrase. *9th European Conference on Speech Communication and Technology*. 2365–2368. https://doi.org/10.21437/Interspeech.2005-37.

Kohler, Klaus J. 1984. Phonetic explanation in phonology: The feature fortis/lenis. *Phonetica* 41(3). 150–174. https://doi.org/10.1159/000261721.

Ladefoged, Peter & Sandra F. Disner. 2012. *Vowels and consonants*. 3rd edition. Malden, MA/Oxford: Wiley-Blackwell. 138–145.

Lisker, Leigh & Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *WORD* 20(3). 384–422. https://doi.org/10.1080/00437956.1964.11659830.

Lisker, Leigh. 1986. "Voicing" in English: A catalogue of acoustic features signalling /b/ versus /p/ in trochees. *Language and Speech* 29(1). 3–11. https://doi.org/10.1177/002383098602900102.

Lisker, Leigh. 2002. The voiceless unaspirated stops of English. In Bruce E. Nevin (edn.), *The Legacy of Zellig Harris*. John Benjamins Publishing Co. 233–240.

Lindsey, Geoff & Péter Szigetvári. 2013–. *CUBE: Current British English searchable transcriptions*. http://cubedictionary.org.

Lombardi, Linda. 1995. Laryngeal neutralization and syllable wellformedness. *Natural Language and Linguistic Theory* 13(1). 39–74.

Mannell, Robert. 2024. *Speech spectra and spectrograms: 4. Some detailed views of "heed."* [PowerPoint.] Macquarie University. https://www.mq.edu.au/__data/assets/pdf_file/0004/911641/heed_closeups.pdf.

Mester, R. Armin & Junko Itô. 1989. Feature predictability and underspecification: Palatal prosody in Japanese Mimetics. *Language* 65(2). 258–293. https://doi.org/10.2307/415333.

Microsoft Corporation. 2024. *Microsoft Excel.* [Computer program.] Version 2406. Build 17726.20126. https://office.microsoft.com/excel.

Muse Group. 2024. *Audacity*. [Computer program.] Version 3.6.1. https://www.audacityteam.org/.

Nádasdy, Ádám. 2008. *Background to English pronunciation: Phonetics, phonology, spelling: For students of English at Hungarian teacher training institutions*. Nemzeti Tankönyvkiadó.

Nance, Claire & Jane Stuart-Smith. 2013. Pre-aspiration and post-aspiration in Scottish Gaelic stop consonants. *Journal of the International Phonetic Association* 43(2). 129–152. https://doi.org/10.1017/s0025100313000042.

OpenAI. 2024. ChatGPT (October 13 version). [GPT-4o model.] https://chatgpt.com/.

Őri, Péter. 2023. *Analyzing laryngeal contrasts*. [Unpublished manuscript.] Budapest: Department of English Linguistics, Eötvös Loránd University. http://seas.elte.hu/w/!analysis/_media/analysis-laryngeal-settext.pdf.

Page, B. Richard. 1997. On the origin of preaspiration in Scandinavian. *American Journal of Germanic Linguistics and Literatures* 9(2). 167–190. https://doi.org/10.1017/s1040820700002857.

Port, Robert F & Michael L. O'Dell. 1985. Neutralization of syllable-final voicing in German. *Journal of Phonetics* 13(1). 455–471.

Ringen, Catherine & Pétur Helgason. 2004. Distinctive [voice] does not imply regressive assimilation: Evidence from Swedish. *IJES* 4(2). 53–71.

Selkirk, Elizabeth O. 1984. On the major class features and syllable theory. In Mark Aronoff & Richard T. Oehrle (edn.), *Language sound structure: Studies in phonology presented to Morris Halle by his teacher and students.* Cambridge: MIT Press. 107–136.

Sievers, Eduard. 1876. *Grundzüge der Lautphysiologie zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen.* Leipzig: Beitkopf und Härtel.

Smith, Caroline L. 2000. Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet. Cambridge University Press. *Phonology* 17(2). 291–295.

Stahlke, Herbert F. W. 2003. Fortis and lenis obstruents in English. *WORD* 54(2). 191–216. https://doi.org/10.1080/00437956.2003.12068832.

Starčević, Attila (2024). Another fortis-lenis language: A reanalysis of Old English obstruents. *Acta Linguistica Academia* (online first). 1–50. https://doi.org/10.1556/2062.2023.00705.

Szigetvári, Péter. 1998. Voice assimilation in Hungarian: The hitches. *The Even Yearbook* 3. 223–236.

Szigetvári, Péter. 2019. Where are syllables? *Radical: A Journal of Phonology* 1(1). 92–125.

Szigetvári, Péter. 2020. Emancipating lenes: A reanalysis of English obstruent clusters. *Acta Linguistica Academica* 67(1). 39–52. https://doi.org/10.1556/2062.2020.00004.

Törkenczy, Mikós. 2015. Introduction: What is phonological analysis? In Zoltán G. Kiss, Ádám Nádasdy, Attila Starčević, Péter Szigetvári, Miklós Törkenczy (edn.), *Phonological analysis: A coursebook*. Budapest: Department of English Linguistics, Eötvös Loránd University. 1–16.

Törkenczy, Mikós. 2024. *Word Stress Part 2*. [PowerPoint.] Budapest: Department of English Linguistics, Eötvös Loránd University. http://seas.elte.hu/w/!analysis/_media/analysis-stress2-slides.pdf.

Trager, George L. & Henry Lee Smith. 1956. *An outline of English structure*. 3rd edition. Norman, Oklahoma: Battenburg Press.

Twaddell, W. Freeman. 1935. On defining the phoneme. *Language* 11(1). 5–62. https://doi.org/10.2307/522070.

Vaux, Bert. 1998. The laryngeal specifications of fricatives. *Linguistic Inquiry* 29(3). 497–511. https://doi.org/10.1162/002438998553833.

Wells, John C. 1982. *Accents of English 2: The British Isles*. Cambridge University Press.

Wells, John C. 1990. Syllabification and allophony. In Susan Ramsaran (edn.), *Studies in the pronunciation of English. A commemorative volume in honour of A. C. Gimson*. London: Routledge. 76–86.

Wells, John C. 2008. *Longman pronunciation dictionary*. 3rd edition. Harlow: Pearson Education Ltd.

Zec, Draga. 1995. Sonority constraints on syllable structure. *Phonology* 12(1). 85–129. https://doi.org/10.1017/s0952675700002396.

APPENDICES

Left click (or Ctrl+left click) on the image/link to open file.

Appendix A

Material & data collection

(Google spreadsheet)

Material & data collection.xlsx

Appendix B

Segment duration, voicing, ratio

(Google spreadsheet)

Segment duration, voicing, ratio.xlsx

Appendix C

Recordings & segmentations

(.wav & .Textgrid)

https://bit.ly/3AFWy8H

https://bit.ly/3Z63Xs8

https://bit.ly/4e7VCIE

Appendix D

Statistical analysis

Kruskal-Wallis & Mann-Whitney

(Python 3 script)

Statistical analysis.py

Appendix E

Visual representation

Density & Raincloud plots

(Python 3 script)

Density & Raincloud.py

Appendix F

Raw statistical values

(Outcome of the “Statistical analysis.py” script)

1. Duration of the preconsonantal vowel

Kruskal-Wallis H-test statistic: 4.7626333316828005
P-value: 0.09242879951439026
Eta-squared (effect size for Kruskal-Wallis): 0.042523511890025005

Mann-Whitney U Test (Item 1 vs Item 2): U-statistic = 498.5, p-value = 0.03130236570522571, effect size (r) = 0.35505698005698005
Mann-Whitney U Test (Item 1 vs Item 3): U-statistic = 686.0, p-value = 0.5785875702890607, effect size (r) = 0.46288798920377866
Mann-Whitney U Test (Item 2 vs Item 3): U-statistic = 821.0, p-value = 0.1397541767901256, effect size (r) = 0.6001461988304093

2. Duration of the postvocalic stop

Kruskal-Wallis H-test statistic: 8.936243600619962
P-value: 0.01146883639920524
Eta-squared (effect size for Kruskal-Wallis): 0.07978788929124966

Mann-Whitney U Test (Item 1 vs Item 2): U-statistic = 727.0, p-value = 0.7948956432843581, effect size (r) = 0.5178062678062678
Mann-Whitney U Test (Item 1 vs Item 3): U-statistic = 506.0, p-value = 0.016817792843577938, effect size (r) = 0.34143049932523617
Mann-Whitney U Test (Item 2 vs Item 3): U-statistic = 430.0, p-value = 0.006098162242795923, effect size (r) = 0.31432748538011696

3. Duration of the intervocalic consonant cluster

Kruskal-Wallis H-test statistic: 15.30814723291815
P-value: 0.0004741088522868405
Eta-squared (effect size for Kruskal-Wallis): 0.13667988600819778

Mann-Whitney U Test (Item 1 vs Item 2): U-statistic = 935.5, p-value = 0.013445357460270166, effect size (r) = 0.6663105413105413
Mann-Whitney U Test (Item 1 vs Item 3): U-statistic = 610.0, p-value = 0.18356926801161944, effect size (r) = 0.4116059379217274
Mann-Whitney U Test (Item 2 vs Item 3): U-statistic = 323.5, p-value = 9.857867603406372e-05, effect size (r) = 0.2364766081871345

4. Duration of the vowel+consonant cluster

Kruskal-Wallis H-test statistic: 4.010346885829363
P-value: 0.1346369418466669
Eta-squared (effect size for Kruskal-Wallis): 0.03580666862347646

Mann-Whitney U Test (Item 1 vs Item 2): U-statistic = 767.0, p-value = 0.4939129234104769, effect size (r) = 0.5462962962962963
Mann-Whitney U Test (Item 1 vs Item 3): U-statistic = 605.0, p-value = 0.1673413366576384, effect size (r) = 0.40823211875843457
Mann-Whitney U Test (Item 2 vs Item 3): U-statistic = 508.0, p-value = 0.057673249503361314, effect size (r) = 0.3713450292397661

5. Ratio of the vowel's duration to the consonant cluster's duration

Kruskal-Wallis H-test statistic: 67.86079206802142
P-value: 1.8374529767627197e-15
Eta-squared (effect size for Kruskal-Wallis): 0.6058999291787627

Mann-Whitney U Test (Item 1 vs Item 2): U-statistic = 392.5, p-value = 0.001047062703082489, effect size (r) = 0.2795584045584046
Mann-Whitney U Test (Item 1 vs Item 3): U-statistic = 1383.0, p-value = 6.204642863652617e-11, effect size (r) = 0.9331983805668016
Mann-Whitney U Test (Item 2 vs Item 3): U-statistic = 1338.0, p-value = 1.5492002234443865e-12, effect size (r) = 0.9780701754385965

6. Total duration of the occurrences

Kruskal-Wallis H-test statistic: 25.979133425245998
P-value: 2.284035524578784e-06
Eta-squared (effect size for Kruskal-Wallis): 0.2319565484396964
Mann-Whitney U Test (Item 1 vs Item 2): U-statistic = 511.5, p-value = 0.04390743553928636, effect size (r) = 0.3643162393162393
Mann-Whitney U Test (Item 1 vs Item 3): U-statistic = 239.0, p-value = 3.223747941129615e-07, effect size (r) = 0.1612685560053981
Mann-Whitney U Test (Item 2 vs Item 3): U-statistic = 420.0, p-value = 0.004375900686272656, effect size (r) = 0.30701754385964913

7. Ratio of the vowel's duration to the total occurrence's duration

Kruskal-Wallis H-test statistic: 18.090833016062025
P-value: 0.00011793033277904457
Eta-squared (effect size for Kruskal-Wallis): 0.16152529478626806

Mann-Whitney U Test (Item 1 vs Item 2): U-statistic = 684.0, p-value = 0.8524914462327876, effect size (r) = 0.48717948717948717
Mann-Whitney U Test (Item 1 vs Item 3): U-statistic = 1098.5, p-value = 0.0002647792791724951, effect size (r) = 0.7412280701754386
Mann-Whitney U Test (Item 2 vs Item 3): U-statistic = 1024.5, p-value = 0.0002289778771265516, effect size (r) = 0.7489035087719298

8. Duration of voicing in the stop consonant

Kruskal-Wallis H-test statistic: 8.798409187966834
P-value: 0.012287109257881708
Eta-squared (effect size for Kruskal-Wallis): 0.10729767302398578

Mann-Whitney U Test (Item 1 vs Item 2): U-statistic = 760.0, p-value = 0.024022677335549872, effect size (r) = 0.6597222222222222
Mann-Whitney U Test (Item 1 vs Item 3): U-statistic = 391.0, p-value = 0.012628413234683835, effect size (r) = 0.7240740740740741
Mann-Whitney U Test (Item 2 vs Item 3): U-statistic = 290.5, p-value = 0.25303579401133114, effect size (r) = 0.6052083333333333

9. Voicing ratio in the stop consonant

Kruskal-Wallis H-test statistic: 16.311934288075786
P-value: 0.00028701756441130944
Eta-squared (effect size for Kruskal-Wallis): 0.19892602790336325

Mann-Whitney U Test (Item 1 vs Item 2): U-statistic = 764.5, p-value = 0.020712008952807653, effect size (r) = 0.6636284722222222
Mann-Whitney U Test (Item 1 vs Item 3): U-statistic = 451.5, p-value = 0.00017970923690707494, effect size (r) = 0.8361111111111111
Mann-Whitney U Test (Item 2 vs Item 3): U-statistic = 344.0, p-value = 0.018080384752535133, effect size (r) = 0.7166666666666667