Universität Zürich[UZH]

Masterarbeit
zur Erlangung des akademischen Grades
Master of Arts
der Philosophischen Fakultät der Universität Zürich

VOT of bilabial and alveolar stops in Zurich German: a sociophonetic study

Verfasserin: Camille Watter
Matrikel-Nr: 10-297-851

Referent: Prof. Dr. Stephan Schmid

Betreuer: Dr. Simon Clematide

Institut für Computerlinguistik

Abgabedatum: 01.12.2021

Abstract

Young Zurich German speakers seem to increase aspiration in fortis plosives compared to the older generation of speakers. Three groups of Zurich German speakers were compared in this first empirical study on aspiration in fortis plosives to investigate a possible sound change. The control group consists of 10 Zurich German speakers over 60 years old. The 20 secondary school students from Zurich city were split into two groups on the basis of a perception experiment conducted by peers. The so-called *mono* group consists of 10 speakers that were perceived to speak a more traditional Zurich German dialect. The remaining 10 speakers in the so-called *multi* group were perceived to speak multiethnolectal Zurich German. *Multiethnolects* are relatively new ways of speaking that have been observed in multiethnic neighborhoods of European cities, including Zurich.

The analysis is based on the *voice onset time (VOT)* measurements of 526 bilabial and alveolar plosives that were extracted from the DiaPix recordings from the corpus of the *Phonetic features of (multi-)ethnic urban vernaculars in German-speaking Switzerland* project, which is currently carried out at the University of Zurich.

The data analysis shows a clear generational difference in the use of aspiration in fortis plosives and points to a gradual sound change. Younger Zurich German speakers clearly produce longer VOT compared to the control group. Whereas the older Zurich German speakers exhibit a clear three-way contrast in plosives [b̥ p p^h], the multiethnolect speakers rather display a two-way contrast [b̥ p^h]. The *mono* group shows a clear three-way contrast only in alveolar stops [d̥ t t^h].

Zusammenfassung

Jüngere Sprecher des Zürichdeutschen zeigen im Vergleich zur älteren Generation eine zunehmende Aspiration in Fortis-Plosiven. In dieser ersten empirischen Studie zur Aspiration in Fortis-Konsonanten wurden drei Sprechergruppen verglichen, um diesem möglichen Lautwandel auf den Grund zu gehen. Die Kontrollgruppe besteht aus zehn Zürchern, die über 60 Jahre alt sind. Die zwanzig jüngeren Zürcher Sekundarschüler wurden aufgrund eines Wahrnehmungsexperiments mit gleichaltrigen Schülern in zwei Gruppen geteilt. Die sogenannte *mono* Gruppe besteht aus zehn Sprechern, deren Zürichdeutsch als traditional eingestuft wurde. Die restlichen zehn Schüler der sogenannten *multi* Gruppe wurden als multiethnolektale Sprecher wahrgenommen. *Multiethnolekte* bezeichnen relative neue Sprechweisen, die in multikulturellen Vierteln in europäischen Städten entstanden sind, und auch in der Stadt Zürich beobachtet worden wird.

Die Analyse basiert auf Messungen der sogenannten *voice onset time (VOT)* von 526 bilabialen und alveolaren Plosiven, die aus dem DiaPix Korpus des Projekts *Phonetische Merkmale von multiethnischen urbanen Sprachvarietäten in der deutschsprachigen Schweiz*, welches gegenwärtig an der Universität Zürich durchgeführt wird, extrahiert wurden.

Die Datenanalyse zeigt einen klaren Unterschied im Gebrauch von Aspiration zwischen den Generationen und weist auf einen graduellen Lautwandel hin. Die jüngeren Sprecher beider Gruppen produzieren deutlich längere VOT Messungen als die ältere Kontrollgruppe. Während die ältere Generation einen deutlichen dreiteiligen Kontrast zwischen [b̥ p p^h] aufweist, zeigen die multiethnolektalen Sprecher einen zweiteiligen Kontrast [b̥ p^h]. Der dreiteilige Kontrast konnte in der *mono* Gruppe nur bei alveolaren Plosiven nachgewiesen werden [d̥ t t^h].

Acknowledgements

First, I want to thank Prof. Dr. Stephan Schmid for coining the idea for this thesis. I would like to express my sincere gratitude for the valuable guidance, dedication, patience, continuous input, and support in every step of the process.
Many thanks also to Dr. Simon Clematide and Dr. phil. Dieter Studer-Joho for valuable instructions and recommendations.

A warm and special thank you for Dr. Marie-Anne Morand who provided technical expertise, helpful feedback, and invaluable suggestions.

Many thanks also to Sandra Schwab whose Praat script that automatically opens and saves TextGrids was a tremendous help for the annotation process.

Last but certainly not least, I want to thank Sebi for the support, providing study music, encouragement throughout the writing process, and proofreading. Finally, a big thank you to Andrea, Sam and Ines for their kind support.

Contents

Abstract i

Acknowledgement iii

Contents iv

List of Figures vii

List of Tables viii

List of Acronyms ix

1 Introduction 1

1.1 Background 1

1.2 Research questions and hypotheses 2

1.3 Thesis structure 3

2 Characteristics of plosives and related work 4

2.1 Voicing, aspiration and VOT 4

2.2 Characteristics of plosives in Zurich German 7

2.3 Ethnolectal Zurich German 8

3 Method 11

3.1 DiaPix recordings 11

3.2 Participants 12

3.3 Data processing and plosive selection 13

3.3.1 Detecting plosives in the DiaPix corpus 14

3.4 Plosive corpus 16

3.4.1 Automatic annotation with Webmaus 18

3.4.2 Manual segmentation in Praat 18

3.5 Annotation guidelines 19

3.5.1 Detailed annotation of plosives 20

3.6 Durational measurements 23

3.7 Statistical analysis 24

4 Results 27

4.1 Results of VOT measurements 28

4.2 Group comparison 31

4.3 Place of articulation . . . 33
4.4 Aspiration . . . 34
4.5 Gender . . . 37
4.6 Speech rate normalization . . . 40

5 Discussion **43**
5.1 Comparison between groups and place of articulation . . . 43
5.2 Differences in [±Aspiration] . . . 44
5.3 Gender . . . 47
5.4 Speech rate normalization . . . 49
5.5 Speaker variability . . . 49
5.6 Limitations . . . 50

6 Conclusion **54**

References **57**

7 Tables **61**

8 Figures **65**
8.1 Quantile-quantile and density plots per comparison . . . 65
8.2 Boxplots of aspiration comparison within group . . . 72
8.3 Boxplots of gender comparison within group . . . 73
8.4 DiaPix picture pairs . . . 74

9 Scripts **75**
9.1 Channel and tier extraction from sound files and textgrids . . . 75
9.2 Find and extract plosives from textgrids . . . 76
9.3 Find and extract plosives from Praat output . . . 78
9.4 Extract annotation and sound files for Webmaus . . . 79
9.5 Alter tiers in textgrids . . . 80
9.5.1 Alter tier of Webmaus output . . . 80
9.5.2 Inserting tier . . . 81
9.5.3 Duplicate segment tier . . . 81
9.5.4 Insert dummy annotation . . . 82
9.6 Cut sound and textgrid at plosive with context . . . 83
9.6.1 Cut all files automatically . . . 83
9.6.2 Save sound and textgrid of manual selection . . . 84
9.7 Measurement calculation . . . 84
9.7.1 Move boundaries to nearest zero crossing . . . 84
9.7.2 Durational measurements . . . 85
9.8 Other utility scripts . . . 86
9.8.1 Sort and count plosives . . . 86
9.8.2 Move files to different folder . . . 87

9.8.3 Collect unique name of each plosive in corpus 87
9.8.4 Rename all plosives by giving each file unique number 88
9.8.5 Save all selected files in Praat object window 88
9.9 R Scripts 89
9.9.1 R script for statistical analysis and visualizations 89
9.9.2 R script pVOT, further analysis, and visualizations 100

List of Figures

2.1 Aspiration differences in alveolar plosives . . . 5
3.1 Original sound and TextGrid file . . . 14
3.2 Input and output of Webmaus . . . 17
3.3 Shortened sound and TextGrid files . . . 20
3.4 Detailed annotation of target plosive . . . 21
3.5 Final TextGrid file with adjusted boundaries . . . 23
4.1 Comparison of VOT of all plosives in terms of [±Aspiration] per group . . . 27
4.2 Quantile-quantile and density plots for VOT in the *mono*, *multi*, and *o60* groups . . . 31
4.3 VOT comparison of all plosives per group . . . 32
4.4 Comparison of VOT durations of bilabial and alveolar plosives . . . 33
4.5 Group comparison of aspirated and unaspirated bilabial and alveolar stops . 35
4.6 Gender differences of in VOT measurements . . . 38
5.1 Speaker variability in VOT per group . . . 51
8.1 Quantile-quantile and density plots for plosive /p/ per group . . . 65
8.2 Quantile-quantile and density plots for plosive /t/ per group . . . 65
8.3 Quantile-quantile and density plots for aspiration [p^h] per group . . . 66
8.4 Quantile-quantile and density plots for aspiration [p] per group . . . 66
8.5 Quantile-quantile and density plots for aspiration [t^h] per group . . . 67
8.6 Quantile-quantile and density plots for aspiration [t] per group . . . 67
8.7 Quantile-quantile and density plots for gender /p/ f per group . . . 68
8.8 Quantile-quantile and density plots for gender /p/ m per group . . . 68
8.9 Quantile-quantile and density plots for gender /t/ f per group . . . 69
8.10 Quantile-quantile and density plots for gender /t/ m per group . . . 69
8.11 Quantile-quantile and density plots for pVOT [p^h] per group . . . 70
8.12 Quantile-quantile and density plots for VOT [p^h] per group . . . 70
8.13 Quantile-quantile and density plots for pVOT [t] per group . . . 71
8.14 Quantile-quantile and density plots for VOT [t] per group . . . 71
8.15 Comparison of aspiration per PoA for *mono* . . . 72
8.16 Comparison of aspiration per PoA for *multi* . . . 72
8.17 Comparison of aspiration per PoA for *o60* . . . 72
8.18 Comparison of gender for *mono* . . . 73
8.19 Comparison of gender for *multi* . . . 73
8.20 Comparison of gender for *o60* . . . 73
8.21 DiaPix used to elicit spontaneous speech by participants . . . 74

List of Tables

2.1 Types of plosives 6
3.1 Plosive count per speaker and group 13
3.2 Overview of number of plosives in the analyzed corpus in terms of aspiration, stress, loan words, and syllable count 16
3.3 Sequence of statistical tests 25
4.1 VOT measurements and token count per group 28
4.2 Plosive count, VOT measurements, and tests for normal distribution and homogeneity of variance per compared factor. 29
4.3 Results of non-parametric ANOVA show if there is a significant difference in VOT between groups. 29
4.4 Results of post hoc test identify groups that are significantly different from one another. 30
4.5 Results of pairwise comparison tests if VOT is produced significantly different in aspirated and unaspirated stops within groups. 30
4.6 Comparison of absolute (VOT) and normalized (pVOT) voice onset time . 41
4.7 ANOVA and post hoc test of absolute (VOT) and normalized (pVOT) voice onset time 41
7.1 Overview of all 106 types of target words and occurrences in corpus 61
7.2 pVOT and VOT comparison of two syllable words 62
7.3 Statistical test to identify groups that are significantly different for two syllable words 62
7.5 Test corpus 63
7.4 Speaker variability per group 64

List of Acronyms

IPA	International phonetic alphabet
MAD	Median absolute deviation
MoA	Manner of articulation
ms	Milliseconds
PoA	Place of articulation
pVOT	Normalized voice onset time
SD	Standard deviation
VOT	Voice onset time

1 Introduction

1.1 Background

Zurich German is a one of many Swiss German dialects of Switzerland that is spoken in Zurich city and roughly within the boarders of the canton of Zurich [Fleischer and Schmid, 2006, 243]. Weber [1948] suggests that there are several Zurich German dialects, in this study, however, Zurich German will not be differentiated into subcategories and varieties, and is treated as a uniform dialect.

The characteristics of plosives in Zurich German are distinct from those of other languages. French, for example, distinguishes like many other languages the plosives /b d ɡ/ and /p t k/ with the same respective place of articulation by the use of voicing, i.e. vocal fold vibration. Whereas the bilabial plosive with voicing results in /b/, the voiceless production in the same place of articulation results in /p/. The difference between the two plosives in the example can be measured using the voice onset time (VOT) [Lisker and Abramson, 1964]. In Zurich German however, all plosives are voiceless. The contrast between homorganic stop consonants is differentiated primarily by the length of closure duration. A stop consonant with a short closure duration is called *lenis* (/b̥ d̥ ɡ̊/) and with a long closure duration *fortis* (/p t k/) [Ladd and Schmid, 2018].

Another phonetic characteristic of Swiss German stops is the lack of aspiration [Fulop, 1994, 59]. There are, however, exceptions like loan words from Standard German, where aspiration is determined lexically. As a result, fortis plosives are aspirated before a stressed vowel as in [t^{h}eː] *Tee* ‘tea’ [Fleischer and Schmid, 2006, 244]. Aspiration may also appear in loan words from English, e.g. [p^{h}ɑːrti] *Party*, as well as some proper names as in *Peter*, but not in *Thomas* [Schmid, 2019a, 15]. Differences in aspiration can be measured with the voice onset time.

The canton of Zurich has around 1.55 million inhabitants[1], of which almost a third live in the city of Zurich[2]. A variety of languages are spoken in Zurich due to immigration. In consequence, the actual number of Zurich German speakers is certainly lower than the number of inhabitants [Fleischer and Schmid, 2006, 243]. Other languages, besides the national languages, that are frequently spoken in the city of Zurich include among others

[1] `https://www.zh.ch/de/soziales/bevoelkerungszahlen.html?keyword=einwohner#/home` Accessed: 02.11.2021.

[2] `https://www.stadt-zuerich.ch/prd/de/index/statistik/publikationen-angebote/publikationen/webartikel/2021-02-18_Die-Stadtzuercher-Bevoelkerung-im-Jahr-2020.html` Accessed: 02.11.2021.

English, Serbian, Croatian, or Spanish[3].

Since the millennium, new vernaculars spoken by adolescents have been observed in multiethnic districts in multiple European cities. Such a new manner of speaking called *multiethnolect* or *multiethnolectal speech* has also emerged in multicultural neighborhoods in Zurich [Schmid, 2010; Morand et al., 2019, 2021]. Ethnolectal speech differs syntactically, lexically, and phonetically from traditional Swiss German dialects [Morand et al., 2019].

1.2 Research questions and hypotheses

The use of aspiration in traditionally unaspirated stops of Swiss German has increased among younger generations [Schifferle, 2010, 11; Ladd and Schmid, 2018, 232, 246; Leeman et al., 2020, 64]. However, there has been no empirical study specifically dedicated to aspiration in fortis stops in Zurich German so far. The generational difference in plosives in terms of aspiration may indicate a sound change in progress. We will therefore apply the co-called *apparent time paradigm* analysis which compares data of older and younger speakers of the same language community that are collected at the same point in time in order to detect a possible change of a certain phonetic phenomenon. The paradigm assumes that the language of the older generation of speakers remains unchanged since their adolescence and therefore represents the language before a sound change [Labov, 1994, 43-72; Kleber, 2016, 132-133].

The study focuses on the use of aspiration in bilabial /p/ and alveolar /t/ fortis plosives. The velar fortis stop consonant /k/ is excluded because it is far less common a in word-initial position in the Zurich German dialect compared to Standard German. Word-initial lexical velar stops /k/ in Standard German as in *Kiste* /ˈkɪstə/ 'box' are typically realized as fricatives in Zurich German /ˈxiʃtə/.

The study investigates the voice onset time in plosives in Zurich German in three different groups of speakers. One group consists of native speakers over 60 years old, whereas the other two comprise of secondary school students who speak either traditional or multiethnolectal Zurich German. The categorization of the adolescent Zurich German speakers are based on a sociolinguistic perception experiment conducted with teenage Zurich German speakers of a third secondary school. These pupils rated recordings of the Zurich German speakers according to their perception of how multiethnolect the recorded speakers sounded [Morand et al., 2020b].

Ladd and Schmid [2018, 246-247] observe an increase in aspiration of bilabial and alveolar stop consonants by younger Zurich German speakers especially in new lexical items such as *Panda* or *Porsche*. In contrast, older Zurich German speakers do not aspirate the same words. They conclude that the sound change in terms of aspiration probably proceeds

[3] `https://www.stadt-zuerich.ch/prd/de/index/statistik/publikationen-angebote/publikationen/webartikel/2012-09-06_Wie-spricht-Zuerich.html` Accessed: 02.11.2021.

through lexical diffusion.

On the other hand, Morand et al. [2021, 14-15] report an increase of aspiration in word-initial bilabial and alveolar fortis plosives by multiethnolect speakers. They argue, however, that it remains unclear if this is a specific feature of multiethnolectal speech or if this suggests a starting sound change. They further report frequent aspirated stops in younger subjects that are not perceived as multiethnolectal Zurich German speakers. These findings could indicate a more gradual sound change in the Zurich German dialect, where traditionally unaspirated stops are being produced with aspiration.

This study aims to primarily explore the hypothesis of a gradual sound change by comparing bilabial and alveolar stops produced by older as well as younger traditional and multiethnolect Zurich German speakers. We expect that younger Zurich German speakers of both groups tend to produce bilabial and alveolar plosives with more aspiration than older Zurich German speakers. Older native speakers should therefore display less aspiration, i.e. shorter VOT, compared to the adolescent speakers. We additionally expect a difference between the two groups of younger speakers and hypothesize that the multiethnolect speakers should display the longest voice onset time of all three groups.

We further differentiate between traditionally aspirated and unaspirated plosives within each group. We expect that older Zurich German speakers produce aspirated and unaspirated stops distinctly different. Moreover, we want to investigate if adolescent speakers make a significant contrast between aspirated and unaspirated stops.

Besides age, aspiration, and multiethnolect, the factor gender is taken into account to investigate if there are durational differences in VOT production by female and male speakers in each individual group.

1.3 Thesis structure

The following chapter first explains the general characteristics of stop consonants and points out the unique traits of plosives in Zurich German. The second chapter additionally discusses the term *multiethnolect* and introduces the SNF research project *Phonetic features of (multi-)ethnic urban vernaculars in German-speaking Switzerland* and the connection to this study. Chapter 3 describes the participants, the DiaPix corpus, as well as the method of finding and processing the fortis plosives in the DiaPix corpus using Praat and Python scripts as well as the web application Webmaus. The third chapter further describes the corpus of 526 fortis plosives and outlines the annotation guidelines of the recordings, as well as the durational measurements. Lastly, chapter 3 illustrates the comparisons that are made between the three groups and lists the applied statistical calculations using R. The fourth chapter reports the results of the statistical analysis. Chapter 5 discusses the findings and the final sixth chapter sums up the conclusions from the study.

2 Characteristics of plosives and related work

2.1 Voicing, aspiration and VOT

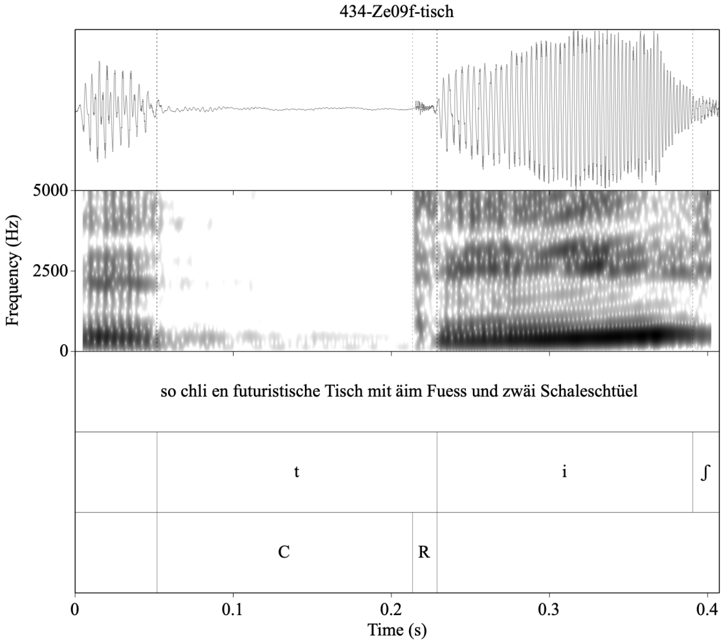
Typically, we produce speech sounds by pushing air from the lungs through the vocal folds in the larynx. A voiced speech sound is created when the airstream causes the vocal folds to vibrate. The waveform is the visual representation and shows a periodic pattern, as shown in the vowel [i] in the top section of Figure 2.1b. The sound can be modified in the oral cavity by obstructing the airstream with articulators such as lips, teeth, or the position of the tongue. For example, a wide open mouth causes almost no obstruction to the airstream which produces the vowel [a].

When the vocal folds are spread open, they do not vibrate. If simultaneously the airstream is exhaled through a heavily obstructed oral tract, such that a turbulent airstream is created, noise instead of a sound is produced. This results in a highly irregular waveform with no periodic pattern, as shown in the interval labeled R in the top section of Figure 2.1b. For instance, placing the upper teeth on the lower lips where the air is forced though the narrow opening creates the labiodental fricative [f].

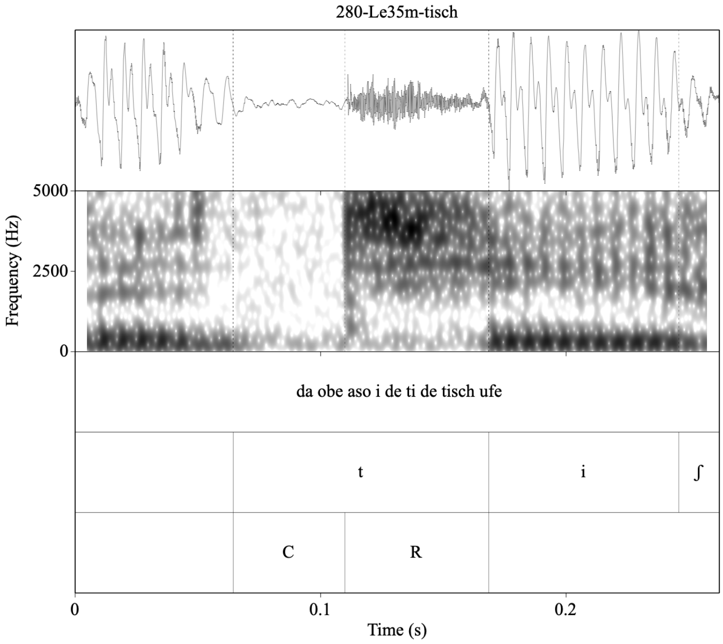
A plosive consists of the two phases *closure* and *release (VOT)* which are annotated with C and R in Figure 2.1. The top part of each figure displays the waveform which is defined by the amplitude over time. The lower graphs show the spectrogram that displays changes in spectral frequencies over time. The degree of shading within the spectrogram indicates the relative intensity of difference spectral bands. A darker area signifies higher spectral energy.

Every plosive starts with a total closure of the oral cavity that prohibits the constant airstream coming from the lungs from exiting. For example, the oral tract is sealed by pressing the lips together when producing the bilabial stop /p/. As a consequence, pressure builds up within the closed oral tract. Since there is no air exiting the mouth, this part of the stop consonant is silent, as demonstrated in the interval labeled C in Figure 2.1a.

The burst of a plosive is at the same time the offset of the closure phase and the onset of the release. At this moment, the oral cavity is opened which causes a sudden audible release of the pressured air. The burst is clearly visible in Figures 2.1a and 2.1b in the spectrogram as dark shading over a large frequency range after the silent period.



(a) *tisch* without aspiration [t]



(b) *tisch* with aspiration [t^h]

Figure 2.1: Aspiration differences in alveolar plosives

During the release phase, the remaining airstream from the lungs exits the oral cavity which is called *aspiration*. It is visualized as noise in the spectrogram and as an irregular pattern in the waveform in Figure 2.1b in the interval labeled *R*. The interval between the burst and the onset of vibration of the vocal folds of the following vowel is called the *voice onset time* (VOT), i.e. the release. The vibration of the vocal folds or *voicing* is a 'periodic pulsing at the frequency of the voice pitch' [Lisker and Abramson, 1964, 387]. Besides the regular pattern in the waveform, voicing is indicated as the dark bar at the bottom of the spectrogram with regular vertical columns that represent the regular pattern of harmonics caused by the vibrating vocal folds. The VOT can differ greatly between speakers in the same word. Figure 2.1a shows an unaspirated alveolar with a VOT of 16 ms by the participant *Ze09* from the group of older Zurich German speakers. The closure measures 162 ms. In comparison, Figure 2.1b displays an aspirated alveolar stop with a VOT that lasts nearly three times as long with 59 ms. The closure duration in the second example measures only 46 ms, and also shows some mild vibration that is visible in both the regular waveform as well as in the light grey voice bar at the bottom of the spectrogram. This might be caused by reverberation from the preceding vowel.

Lisker and Abramson [1964, 387] claim that only one feature at a time – voicing or aspiration – tends to be present in the spectrogram. So if either voicing or aspiration is more prominently displayed in the spectrogram the other is absent or scarcely visible. Additionally, the features are easily distinguishable due to the either regular (voicing) or irregular (aspiration) patterns in both the wave form and the spectrogram. This facilitates the identification and annotation as well as the measurement of intervals such as the VOT.

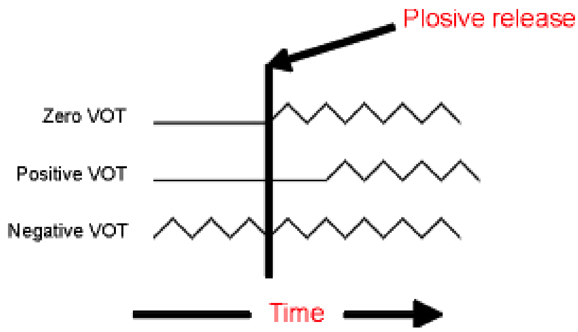


Figure 2.2: Differences in voice onset time

Plosive	Description	VOT
[pa]	voiceless unaspirated plosive	0 - 30ms
[p^{h}a]	voiceless aspirated plosive	> 30ms
[ba]	voiced plosive	negative

Table 2.1: Types of plosives

There are three types of stop consonants that can be differentiated by the voice onset time as shown in Figure 2.2[1] and in Table 2.1 [Schmid, 2020, 12].

The VOT is relatively short in a voiceless unaspirated plosive [pa] where the vocal folds start vibrating approximately at the burst or shortly after. Voicing is represented by the zigzag line in Figure 2.2. In contrast, there is a significant delay in voicing after the burst in a voiceless aspirated plosive [p^{h}a]. And lastly, a negative VOT means continuously vibrating vocal folds in the closure and release phase.

Most languages include two of the three types of plosives. German for example is a so-called 'aspirating' language and distinguishes between a short-lag [pa] and long-lag [p^{h}a] VOT which phonologically translates to /ba/ ∼ /pa/. On the other hand, French is

[1] `https://www.phon.ucl.ac.uk/home/johnm/siphtra/plostut2/plostut2-2.htm` Accessed: 15.09.2021.

considered to be a 'true voice' language which differentiates between negative [ba] and zero [pa] VOT where the phonetic output is the same as the underlying phonological structure /ba/ ∼ /pa/ [Beckman et al., 2013, 260-268].

2.2 Characteristics of plosives in Zurich German

Zurich German has six stop consonants /p b̥ t d̥ k g̊/ that have two shared characteristics: All are voiceless and unaspirated [Fleischer and Schmid, 2006]. There are however some exceptions that are indeed aspirated such as [ˈpʰiksəl] *Pixel.* Aspirated stops are found in loan words from mostly English and Standard German as well as in some proper names, as Zurich German speakers clearly aspirate the bilabial stop in [ˈpʰɑːrti] *party* but do not aspirate the plosive in [ˈpit͡sɑ] *pizza* [Würth, 2002, 7]. The aspiration of plosives in Swiss German is by no means a modern phenomenon and has been reported in dialectological literature and glossaries since the late 19th century. As previously mentioned, there seems to be increase in aspiration in stop consonants of younger speakers in Swiss German, as well as a certain amount of variability between speakers [Schifferle, 2010]. This inevitably leads to variably produced words. However, speakers are generally aware of the distinction between words that are produced with or without aspiration. The production of aspiration is therefore lexically determined and broadly consistent across speakers [Ladd and Schmid, 2018, 232].

Schifferle [2010, 44-45] conducted an empirical study with 8 students from the cantons Zurich and Aargau and found highly varying VOT measurements. In a comparison of bilabial stops using the word *Puls* 'pulse', he reported a VOT of 18 milliseconds as well as 56 milliseconds. The comparison of alveolar stops in *Takt* 'beat' shows similar results as he measured a VOT of 20 milliseconds in the unaspirated production and 62 milliseconds in the aspirated version. This suggests that Zurich German phonetically distinguishes three types of plosives [b̥ p pʰ]. Ladd and Schmid [2018, 237] report further evidence for a three-way stop contrast in phonetic categories and found a clear distinction between the three types lenis, fortis and aspirated fortis plosives in both closure and release duration. Phonologically, however, there are no set of rules for aspiration in stop consonants, but there are possible minimal pairs such as in [tæi̯l] 'component' and [tʰæi̯l] 'object' [Schmid, 2019b]. The contrast is however 'marginal' as described by Renwick and Ladd [2016].

The contrast of homorganic stops in Zurich German is called *fortis* (/p t k/) and *lenis* (/b̥ d̥ g̊/). The terms were introduced by Winteler [1876, 21-23] who recognized that Swiss German obstruents are neither contrasted by aspiration nor voicing. They show however a difference in intensity, where fortis plosives have higher intensity compared to lenis stops. He further states that Swiss German homorganic obstruents should therefore be distinguished by fortis *hart* 'strong' and lenis *weich* 'soft'.

The terms fortis and lenis have been attributed different meanings in the history of research and definitions vary among linguists and phoneticians. The corresponding acoustic correlates have also been widely discussed. However, modern interpretations of the terms are based on Winteler's terminology and findings [Braun, 1988, 6 pp.]. For example, Sievers [1876, 65] supports the terms fortis and lenis by Winteler, and sees a primary difference in intensity on which the duration of a consonant is directly dependent. Lisker and Abramson [1964, 385-386] have a critical stance on the concept of fortis and lenis. They identify three phonetic dimensions to contrast the homorganic plosives: voicing, aspiration and articulatory force which describes the loudness of the burst of a stop consonant, and call the latter the fortis/lenis or tense/lax contrast. They argue on the basis of English stops that the definition of fortis and lenis depends on vague phonetic descriptions and therefore doubt that it is sufficient to physically differentiate the sets of /p t k/ and /b d g/. Thus, they favor voice onset time as distinction of homorganic plosives [Lisker and Abramson, 1964, 420]. In response to these statements, Enstrom and Spörri-Bütler [1981, 138] prove that the distinction of word-initial stops in Zurich German by VOT is questionable. Even though the articulatory features to differentiate homorganic stops have not been reliably specified, they recommend the fortis/lenis opposition introduced by Jakobson et al. [1972] that defines the contrast as a distinctive feature applicable to Swiss German plosives. Fulop [1994, 60] confirms that voice onset time is not suitable to distinguish homorganic stops in Swiss German by showing that the VOT is identical for both fortis and lenis plosives. Moreover, he showed differences between plosives in word-internal and word-final positions where the closure of fortis stops lasted four times longer than the closure duration of lenis stops. Willi [1995, 261; 1996, 195 pp.] found similar results in closure duration differences for stops in word-internal positions. A contemporary approach suggests that the fortis/lenis opposition in Swiss German is a contrast between geminates and singletons [Kraehenmann, 2001; Würth, 2020]. Fleischer and Schmid [2006, 245] refer to the terms fortis and lenis as phonologically distinct homorganic unvoiced stop consonants that are distinguished by an acoustic correlate other than voice.

Both fortis and lenis stops occur in all positions of a word: word-initial, word internal, and word-final [Fleischer and Schmid, 2006]. However, the velar fortis plosive /k/ is rare in word-initial position in Zurich German. Words with a word-initial voiceless velar stop in Standard German are produced with a fricative /x/ or affricate /k͡x/, e.g. [xɒltː] *kalt* 'cold' and [k͡xlɒːr] *klar* 'clear' in Zurich German [Würth, 2020, 24]. Würth [2002, 12] even states that the word-initial velar fortis plosive /k/ does not exist in the native Zurich German vocabulary. However, there are some rare examples with a word-itinital velar fortis plosives such as in the past participle [kɛː] of the verb *geben* 'give'. In consequence, the velar stops are not part of this study.

2.3 Ethnolectal Zurich German

So-called *multiethnolectal speech* has been observed since the late 1980s in Germanic languages in many urban areas in Europe. The way of speaking emerges from neighborhoods with a high percentage of immigrants and serves as an expression of group identity

[Morand et al., 2019]. One of the first research projects on multiethnolectal speech was performed by Rampton [1987]. It analyzes the use of English in multilingual peer groups from a sociolinguistic perspective by discussing identity, social power, and marginalization. Further research on the multiethnolectal effect on British English would follow, notably by Cheshire et al. [2008] who discuss multiethnolectal speech of multiethnic peer groups in London from a phonetic and morphosyntactic perspective. Multiethnolectal speakers were also widely studied in Scandinavian languages (among others by Quist and Svendsen [2010]), and French (for example by Fagyal and Torgersen [2018]). Auer [2003, 258-9] investigated morphosyntactic features of multiethnolectal German and identified, among others, the change of gender of nouns, the omission of prepositions as well as the omission of definite and indefinite articles as the main characteristics of multiethnolectal German.

In the last two decades, multiethnolectal speech has also been observed in different Swiss German dialects. Schmid [2010] discusses the three types of ethnolect proposed by Auer [2003] in Zurich German and describes the speech properties of multiethnolectal Zurich German speakers. Tissot et al. [2011, 322] introduce the term *ethnolektales Schweizerdeutsch* 'ethnolectal Swiss German' that in short describes stylistic variations of Swiss German dialects used by speakers of all generations in multilingual, multiethnic, and predominantly urban contexts. They further present a first list of morphosyntactic, phonetic, and prosodic characteristics of ethnolectal Swiss German. They observe change of gender in nouns, omission of articles, omission of prepositions, omission of impersonal *es* 'it', voicing of lenis consonants, as well as the absence of Sandhi (assimilation of consonants across word boundaries typical for Swiss German) as properties of multiethnolectal speech [Tissot et al., 2011, 341]. The empirical study by Bruno [2019] confirms the omission of definite and indefinite articles as well as the use of deviating article forms for multiethnolectal Zurich German.

This study is linked to the project *Phonetic features of (multi-)ethnic urban vernaculars in German-speaking Switzerland* at the University of Zurich[2] which will be referred to as the *Ethnolect* project in this study. The *Ethnolect* project is an extensive study on multiethnolectal Zurich German and analyzes the phonetic features of multicultural speech of adolescent speakers in Zurich German on both the segmental and the suprasegmental level.

Among other features, the project includes a study on the voicing of traditionally unvoiced lenis plosives in Zurich German. Morand et al. [2019] showed that both multicultural adolescent speakers and monocultural young adults voice lenis plosives. However, there is a signifiant difference between the two groups. The monocultural speakers show a more traditional fortis-lenis pattern in plosives. The occasional voicing of lenis stops can be explained by inter speaker variation and context assimilation. The multiethnolects demonstrate a rather consistent voicing pattern in lenis plosives. Many native languages of the multiethnolect speakers have a voice contrast in stop consonants. Therefore, the

[2] `https://www.cl.uzh.ch/de/phonetics/forschung/Sociophonetics-and-dialectology/phonetic-features_MEZ.html` Accessed: 03.09.2021.

speakers might transfer the voicing pattern from their native language to the articulation of Zurich German lenis plosives.

The *Ethnolect* project additionally includes a study on voiced lenis plosives and their use by multiethnolects as a sociophonetic marker. Morand et al. [2020a] analyze the speech rhythm of multiethnolectal speech and argue that the speech rhythm of multiethnolectal Zurich German, the so-called *staccato* rhythm, is syllable-timed and independent of the mother tongue of multiethnolectal speakers. Another study includes sociolinguistic experiments where Morand et al. [2020b] collect and discuss labels attributed to multiethnolectal Zurich German. They additionally conduct a perception experiment that shows that traditional Zurich German and multiethnolectal Zurich German are not distinctly separated, but rather a continuum ranging from one speaking style to the other.

3 Method

This chapter describes the recordings, participants, as well as the corpus. The analysis of this study is based on data provided by the *Ethnolect* project that collected an extensive database including different speaking styles. The chapter further depicts the method of corpus processing, the annotation guidelines, and the statistical analysis.

3.1 DiaPix recordings

The DiaPix data of the *Ethnolect* project is composed of recordings from 15 dyads of speakers who had to solve a common task in conversation based on the DiaPixUK materials [Baker and Hazan, 2011]. DiaPixUK provides twelve picture pairs aimed to elicit interactive spontaneous dialogue between two speakers for phonetic and linguistic analysis. The picture pair consists of two slightly different versions of a cartoon-like picture. The differences between the two versions can either be an omission or an alteration (e.g. of color) of a certain object. Each participant is given one version of the picture pair that is not shown to the conversation partner. The participants then discuss their pictures to find the differences. This method provides a more controlled setting instead of letting participants spontaneously converse about a specific topic. Due to the reference points in the picture pairs, the participants are more likely to produce similar vocabulary or specific keywords which in turn facilitate acoustic phonetic analysis. Furthermore, Baker and Hazan [2011] have shown that both participants contribute equally to the conversation. As a result, there is no bias towards a certain speaker in a recording.

The provided picture pairs by DiaPixUK can easily be altered to suit a specific elicitation topic. The three picture pairs used for the *Ethnolect* data collection are displayed in Figure 8.21 in the appendix. There are about three differences in each quarter designed to elicit word-initial fricatives as well as lenis plosives in Zurich German. Both groups with younger speakers were recorded late 2018 as well as in the beginning of 2019. The recording of the *o60* group took place in the beginning of 2020. The recordings last between 13 and 26 minutes where two participants at a time discuss two picture pairs.

3.2 Participants

The project's corpus includes recordings of 66 participants. There are ten participants around 60 years of age. The remaining participants are teenagers between thirteen and sixteen years old. The first two letters of each speaker's alias refer to the place of recording. *Bu* stands for *Sekundarschule Buhnrain* and *Le* for *Sekundarschule Letzi* which are both secondary schools in the city of Zurich. Lastly, *Ze* represents *Zentrum* which refers to the Phonetics Laboratory of the University of Zurich. The letters *f* and *m* in each alias stand for female and male. The metadata and additional information such as places of residents or the languages spoken by the participants as well as their competence levels are collected in a detailed list which was provided by Marie-Anne Morand.

The selection of speakers of both younger groups was made with the help of Marie-Anne Morand and is based on a so-called *screening score* [Morand et al., 2020b]. A recording snippet of five to seven seconds of each teenager was rated on a scale from 1 to 7 by 40 pupils of the secondary school *Rebhügel* in Zurich. The listeners were asked to rate the speakers on a scale of 1 to 7 according to how ethnolectal their speech sounded. The higher the score of a speaker the more ethnolectal he or she was perceived. The mean value of all listeners' ratings resulted in a screening score for each speaker. Consequently, the 10 speakers with the highest screening scores were selected for the multiethnolect group and accordingly 10 speakers with a low screening score were selected for the younger monolingual speaker group. In summary, the thirty selected speakers are divided in the following three groups that are defined by age and screening score.

- *mono*: 10 teenagers who have been perceived to speak traditional Zurich German according to low screening scores.
- *multi*: 10 teenagers who have been perceived to speak ethnolectal Zurich German according to high screening scores.
- *o60*: 10 adults over 60 years who speak Zurich German as their first language.

The teenagers of the *mono* group are between fourteen and fifteen years old. All started to speak Swiss German before the age of three and never lived outside of Zurich.

The age range in the *multi* group lies between thirteen and sixteen. Three participants started speaking Swiss German before the age of three, whereas the remaining seven participants started speaking Swiss German between the ages three and eleven. Languages spoken before kindergarten by the speakers include Albanian, Arabic, Bosnian, Croatian, Serbian, Singhalese, Spanish, and Turkish.

The 10 participants in the *o60* group are between 62 and 72 years old and started speaking Swiss German before the age of three. *Ze12f* is the only participant that indicated French as the first language but started to speak Swiss German at the age of five. All speakers lived the majority of their youth and life in the canton of Zurich.

mono				multi				o60			
Speaker	/p/	/t/	Total	Speaker	/p/	/t/	Total	Speaker	/p/	/t/	Total
Le02f	6	9	15	Bu02m	3	7	10	Ze03f	5	13	18
Le05f	3	8	11	Bu13m	5	14	19	Ze04m	3	20	23
Le07f	3	15	18	Bu17f	3	6	9	Ze05m	3	21	24
Le08m	5	11	16	Bu18m	3	8	11	Ze06m	2	29	31
Le09f	6	16	22	Bu22f	9	5	14	Ze07f	3	11	14
Le10f	5	5	10	Le22f	3	4	7	Ze08f	2	13	15
Le11f	5	20	25	Le25m	1	17	18	Ze09f	7	41	48
Le18f	3	8	11	Le29m	3	8	11	Ze10f	5	23	28
Le19m	10	18	28	Le31f	4	9	13	Ze11m	1	20	21
Le33f	2	5	7	Le35m	4	4	8	Ze12f	2	19	21
	48	115	163		38	82	120		33	210	243

Table 3.1: Plosive count per speaker and group

Table 3.1 displays the three groups and the corresponding alias of the selected participants, as well as the number of tokens with a word-initial stop /p/ or /t/ produced by each speaker. The analyzed corpus and process of retrieving words with word-initial bilabial and alveolar stop consonants is described in detail in Sections 3.3.1 and 3.4. The *mono* group consists of two males and eight females, the *multi* group includes six males and four females and lastly, the *o60* group contains four males and six females. In summary, there are twelve male and eighteen female participants and consequently there is a notable bias towards female voices in the corpus.

There are considerably more alveolar stops with 407 tokens compared to 119 bilabial stops in the analyzed data. Each group has a comparatively similar amount of bilabial plosives. Though, there are notable size differences with the alveolar plosives per group with 210 alveolar stops /t/ in the *o60* group, contrasting the relative small number of 82 tokens in the multiethnolect group. The *mono* group has 115 tokens with word-initial alveolar stops. Additionally, nearly half of all tokens, 243 of 526 stops, are produced by the *o60* group, and only about a fifth with 120 tokens by the *multi* group. The *mono* group has a total token count of 163.

3.3 Data processing and plosive selection

The corpus was processed in multiple steps using Praat [Boersma and Van Heuven, 2001] and Python [Van Rossum and Drake Jr, 1995] scripts as well as the web application Webmaus[1] [Kisler et al., 2012]. The statistical analysis and data visualization was made with R [R Core Development Team, 2016]. The scripts are collected in Section 9 in the appendix.

[1] `https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface/Pipeline` Accessed: 02.06.2021.

3.3.1 Detecting plosives in the DiaPix corpus

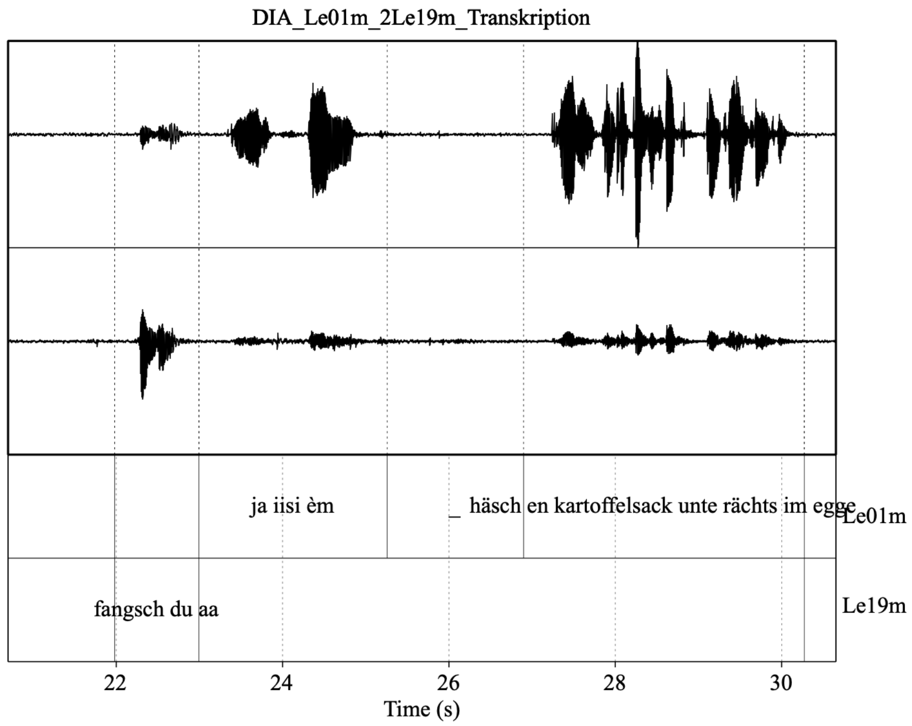


Figure 3.1: Original sound and TextGrid file

Each conversation recording consists of a sound file with two channels, one for each of the two speakers who were recorded with separate microphones. The recordings in the school were made in different rooms that were not set up for proper sound recording. Therefore, the audio of the recordings contains background noise of everyday school life and reverberation. As shown in Figure 3.1, the *Ethnolect* project provided a Praat TextGrid file for each stereo sound file with separate annotation tiers for each of the two speakers. The TextGrid files contain annotations where one utterance of a student equals one annotated unit that ends when the conversation partner starts speaking.

The first Praat script in Section 9.1 extracts each of the two sound channels from the original stereo WAV file and saves each channel in new sound files separately. In the same way, the script splits the annotation tiers into two TextGrid files for each speaker. The top waveform in Figure 3.1 corresponds to the first annotation tier titled *Le01m*, and the second waveform refers accordingly to the tier of speaker *Le19m*. Both the newly created audio and the annotation file are renamed automatically after each speaker with their corresponding alias. As a result, the files of the 30 previously chosen participants could easily be separated from the discarded speakers, as for example *Le01m* who is not part of this analysis.

Notably, the main focus of the *Ethnolect* project does not lie on fortis plosives. The picture pairs are customized to elicit word-initial fricatives as well as lenis plosives and not the needed fortis plosives for this study. The speaking style of the corpus is spontaneous conversation which poses challenges that are not encountered in a controlled speaking style like read speech. There is laughter, coughing, hesitation, or unfinished words and sentences. Moreover, there is mumbling or incomprehensible speech in the audio, as well as reverberation. In consequence, it was not possible to anticipate at this stage of the study how many fortis plosives are contained in the recordings and if there would be enough plosives for the intended analysis. The goal was to obtain 3 to 10 usable fortis bilabial /p/ and alveolar /t/ plosives per speaker. Preferably, all plosives for the analysis should be positioned in the same context: word-initial and preceding a stressed vowel.

Since we could not be sure to gain enough word-initial plosives, the second Praat script in Section 9.2 scans the annotation of all TextGrids in a folder for the characters 'p' and 't' that are immediately followed by a vowel in any position in a word using regular expression[2] [Aho, 1991]. The script outputs a table that displays the speaker, plosive, and position of the matched plosive within the word[3] for every match, as well as the complete annotation of an utterance and the timestamp within the sound file of the beginning of the utterance. An utterance refers to an annotated unit between two boundaries in the original recording, i.e. a unit of uninterrupted speech by one speaker until the conversation partner starts speaking. The annotated utterances vary greatly in length and range from a single word to multiple sentences.

The application of regular expressions in Praat is somewhat limited. For example, the script was not able to search for the plosives `PpTt` at the same time; instead the script only worked efficiently when searching for `Pp` and `Tt` stops separately. Further, it was not possible to search for patterns and simultaneously output the following vowel or preceding word in the same script which is the reason why the output table contains a placeholder 'x'. In this case, the shortcomings of this script can be neglected, because the focus of this initial search was to determine if the recordings contain enough word-initial plosives, and depending on the result, if word internal plosives should be considered for this analysis. The script matched a total of 1713 stop consonants of which 516 are word-initial plosives, which was determined to be enough for the analysis.

The Python script in Section 9.3 goes through all the utterances that were collected in the previous Praat script. A new table is generated containing the preceding word, as well as the last letter of the preceding word for each word-initial plosive found. The script yielded 25 additional word-initial plosives that were overlooked by the Praat script resulting in a total of 541 word-initial plosives. It became apparent that the regular expressions in the Praat script would only match the first occurrence of a 'p' or 't' and ignore the rest of the utterance. For example, the regular expression would only match the first *tüüre* in the utterance *èm bi de pinke huus sind die tüüre händ die tüüre au fänschter* 'uhm at the pink house are the doors have the doors windows too' and move on.

[2] A regular expression, or regex for short, is a user specified pattern of characters according to a syntax that is searched for in an input text. In this implementation, the pattern is designed to identify bilabial and alveolar stops in the annotated utterances of the recordings.

[3] The position of the plosive is noted by a code such as '_PV' that points to a white space, plosive and vowel which suggests a word-initial plosive. The script additionally marks the position with the German terms *Anlaut, Inlaut*, and *Auslaut* 'initial, internal and final sound'.

All plosives and the relevant extracted information were merged manually in a large Google Sheets table. All redundant information was removed, such as multiple cues to the position of a plosive within a word. 15 plosives that were deemed unusable such as wrongly matched plosives *Plakat*, 'poster', exclamations *poa*, or unfinished words *ti* were manually removed from the corpus resulting in a total of 526 bilabial and alveolar plosives for the analysis.

3.4 Plosive corpus

As shown in Table 3.1, the goal of gaining around 3 to 10 plosives was clearly met for the alveolar plosives. However, there are six speakers that do not reach the minimum of 3 bilabial plosives. It was decided to perform the analysis regardless since the analysis prioritizes the comparison per group over individuals.

		mono	multi	o60	Total
Aspiration	+	53	34	23	110
	-	110	86	220	416
Stress on first syllable	+	151	116	235	502
	-	12	4	8	24
Loan word	yes	46	40	24	110
	no	117	80	219	416
Syllable count	1	82	64	97	243
	2	72	45	123	240
	3	8	10	19	37
	4	1	–	4	5
	5	–	1	–	1

Table 3.2: Overview of number of plosives in the analyzed corpus in terms of aspiration, stress, loan words, and syllable count

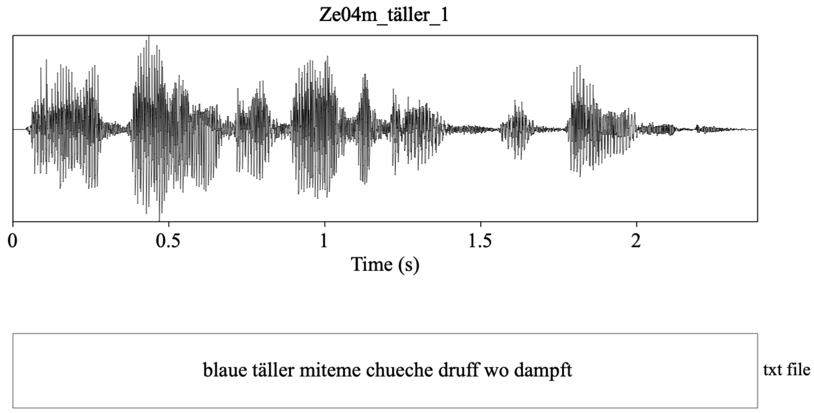
Table 7.1 in the appendix lists all 106 types of words with a word-initial bilabial or alveolar plosive in the analyzed corpus, these are also referred to as *target words*. The table additionally lists for each word type if it is generally aspirated, if it is a loan word, how many syllables it contains, if the first syllable is stressed and how often it occurs in the recordings. The table is sorted by the most frequent target word.

Table 3.2 conveys information on the same categories in terms of token count per group. The corpus for the analysis contains nearly four times as many word-initial unaspirated stops compared to word-initial aspirated plosives. The relatively large amount of traditionally unaspirated plosives provides the basis for the hypothesis analysis of the possible sound change by comparing VOT measurements of plosives between the three groups.

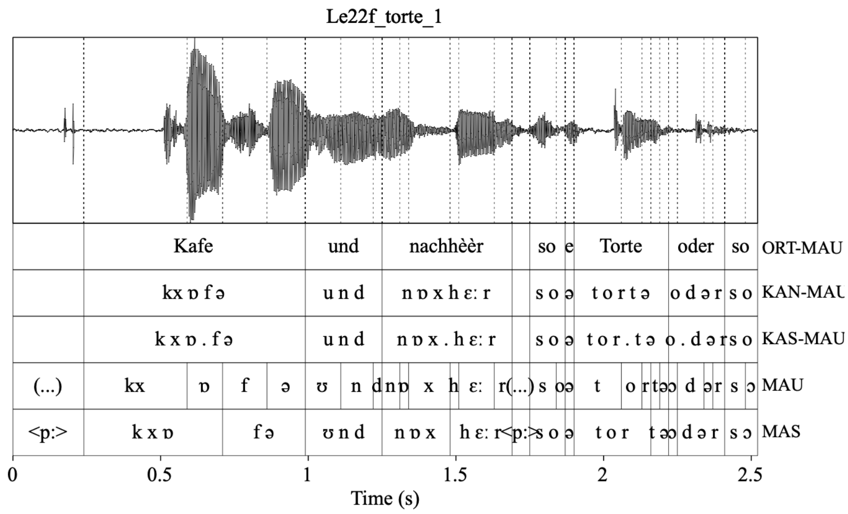
There are only 24 target words that are not stressed on the first syllable. These were still included in the analysis since most of them have a word-initial bilabial stop. By excluding them, some speakers would have little to no bilabial stops.

About a fifth of all tokens are obvious loan words. A target word is categorized as a loan word if it does not appear in the online version of the *Schweizerisches Idiotikon*[4], a lexicography project of Swiss German dialects. Compounds that are not listed in the dictionary and whose components are clearly native are also considered a native word, such as *tüürraame* 'doorframe'.

Additionally, the syllable count of the target words are listed in Tables 3.2 and 7.1. All above mentioned categorizations are based on the perspective of the author of this study.



(a) Sound and text file as input for Webmaus



(b) TextGrid files generated by Webmaus

Figure 3.2: Input and output of Webmaus

[4]`https://www.idiotikon.ch/woerterbuch/idiotikon-digital` Accessed: 15.11.2021.

3.4.1 Automatic annotation with Webmaus

Webmaus[5] is a free web application of the tool MAUS (**M**unich **AU**tomatic **S**egmentation)[6] by the Institute of Phonetics and Speech Processing of the Ludwig Maximilian University of Munich. Webmaus is a service that automatically segments and phonetically annotates speech signals.

Up to this point, there is still one single long sound file per speaker. Instead of annotating the entire sound file, the Praat script in Section 9.4 searches for the target words in all TextGrid files. For each match, the speech signal of the complete utterance is extracted and stored in a separate sound file which is uploaded to Webmaus. Additionally, the annotation of the utterance is extracted and exported into a plain text file (see Figure 3.2a). Because the Webmaus service requires that the sound and text files to be identically named, the Praat script 9.4 renames each newly created sound file after the speaker and the matched plosive. Webmaus unfortunately cannot cope with any non-ASCII characters in the file names. In consequence, file names containing *ä, è, ö* and *ü* were manually renamed before and after the use of Webmaus.

The pipeline `G2P` → `MAUS` → `PHO2SYL`[7] in Webmaus was used to phonetically label the speech signals. Additionally, the following parameters were entered for the service: The chosen language is *German Dieth (CH), Zurich dialect*, the phonetic notation format is IPA, and the output format is defined as a TextGrid file.

3.4.2 Manual segmentation in Praat

The TextGrid files that are generated by Webmaus, as shown in Figure 3.2b, contain five tiers with different automatic annotations of which only the first and fourth are relevant for further processing. The Praat script in Section 9.5.1 renames the first tier containing the word segmentation from the uploaded text file to *words*. Figure 3.3 shows a TextGrid that is further altered but already displays the described tiers that are generated by the script. The phone segmentation and annotation are included in the fourth tier in the TextGrid by Webmaus. The script renames the tier as *segments*. The remaining tiers in the TextGrid files by Webmaus containing phonetic annotation in IPA of word segments, word segments with syllables indicated by points, phone segments, and syllable segments are removed. The script additionally collects the text from all segments in the newly renamed *word* tier in order to create and insert the whole utterance in the *sentence* tier. An additional Praat script in Section 9.5.2 inserts an empty tier named *phases* for the

[5] `https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface/Pipeline` Accessed: 02.06.2021.

[6] `https://www.phonetik.uni-muenchen.de/forschung/aktuelle_projekte/maus.html` Accessed: 15.09.2021.

[7] G2P stands for grapheme-to-phoneme, MAUS for phonetic segmentation, and PHO2SYL for syllabification (phonemic and phonetic) as stated in `https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface/Pipeline` Accessed: 20.09.2021.

manual segmentation of the voice onset time.

At this stage, all sound and TextGrid files still include the time span of the complete utterance (displayed in the tier *sentence*). The files are still unnecessarily large because the measurement of the VOT duration requires only the target word as well as the preceding word for context information. Therefore, the next Praat script in Section 9.6.1 automatically cuts down both the sound and TextGrid files. The script searches for a word-initial plosive using regular expressions. For each match, the script extracts the previous word (if the plosive is not at the beginning of the utterance) and the target word containing the matched plosive. Additionally, the script extracts 0.1 seconds before the preceding word as well as 0.1 seconds after the target word containing the plosive to add enough context and ensure that nothing crucial to the analysis is removed. The script completely relies on the automatic segmentation by Webmaus which unfortunately is not perfectly accurate. A manual inspection revealed that in many cases there was not enough context to determine the precise beginning or ending of relevant segments. In consequence, 59 plosive files of the younger speakers were extracted a second time with 0.4 seconds context before the preceding word and after the target word. Additionally, all files from older Zurich German speakers were extracted again with 0.2 seconds context. There were 31 files that still did not include the necessary segments and needed manual definition of the section to be extracted. This was done with the Praat script in Section 9.6.2 which automatically opens all sound and corresponding TextGrid files in pairs and saves the manually defined section in a new sound and TextGrid file.

The last two Praat scripts in Sections 9.5.3 and 9.5.4 alter the TextGrid files to facilitate manual boundary adjustments before manual annotation. The first script duplicates the tier *segments* and names the duplicate *targetWord* containing the phonetically annotated segments in order to later manually adjust boundaries of the target word containing the word-initial plosive as described in detail in Section 3.5. The second Praat script searches for the start and end point of the target plosive and inserts the found boundaries as dummy intervals C and R in the *phases* tier so that these boundaries do not have to be added manually in all 526 TextGrid files. Instead, this measure accelerates the segmentation process of VOT durations, since only the boundaries need to be adjusted. The dummy *burst* boundary is set right in the middle between the dummy onset and offset boundaries of the target stop consonant by the script.

3.5 Annotation guidelines

In the following study, the notation /p/ with forward slashes implies that all bilabial plosives in the corpus – independent of the categorization [±Aspiration] – are considered. Plosives in square brackets distinguish between the feature aspiration where [p] combine all unaspirated plosives, and [p^h] all aspirated bilabial stops in the corpus according to Table 7.1 in the appendix.

As shown in Figure 3.3, the phonetic annotations and segment boundaries in the tiers *words* and *segments* are not altered from the Webmaus output. Additionally, all boundaries not belonging to the target word were manually removed in the *targetWord* tier. Whereas the phonetic annotations of the target word are not corrected, the segment boundaries are adjusted to facilitate duration calculation of the speaker's speech rate. The boundary adjustment in the *targetWord* tier was primarily based on visual cues in the spectrogram and checked by listening to the segments. These adjustments are not as precise as the segmentation of the plosive phases in the *phases* tier (see Section 3.5.1), but erase extreme mistakes in the automatic annotation by Webmaus. The only phonetic annotation that is corrected is the one of the target stop consonant, as for example the phonetic annotation of the target plosive [t^h] in the tier *targetWordZero* in Figure 3.4.

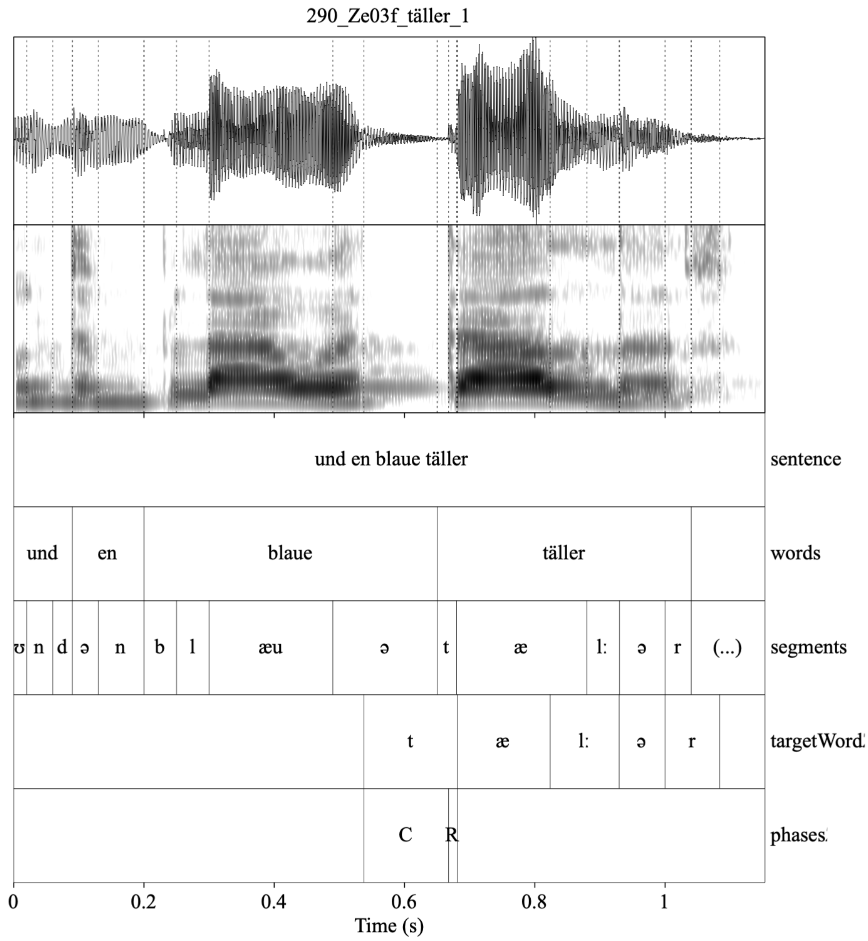


Figure 3.3: Shortened sound and TextGrid files

3.5.1 Detailed annotation of plosives

The annotation of stop consonants primarily follows the waveform of the speech signal. The corresponding spectrogram serves as a reference point but remains secondary to the waveform in the decision of the boundary placement.

There are two phases in each plosive that are annotated: The interval of the closure

C of the oral cavity and the release *R* that starts at the burst and lasts until the following vowel. The resulting three boundaries were annotated in the TextGrid tier *phases* by adapting the following criteria:

- *Onset closure:* The boundary is set at the end of a clear reoccurring pattern in the waveform. The spectrogram shows no clear voice bar at the very bottom and no regular vertical striations that indicate voicing, as shown in Figure 3.4.
- *Burst:* The offset of the closure is simultaneously the burst of the plosive and therefore the onset of the release phase. The boundary is set on the point crossing the zero amplitude line before the sound wave raises to a clear peak in the waveform. The spectrogram displays a dark column over all frequencies.
- *End release:* The boundary is set before a distinct reoccurring periodic pattern of the following vowel in the waveform. The spectrogram displays a clear voice bar as well as second formant which can be automatically displayed by Praat.

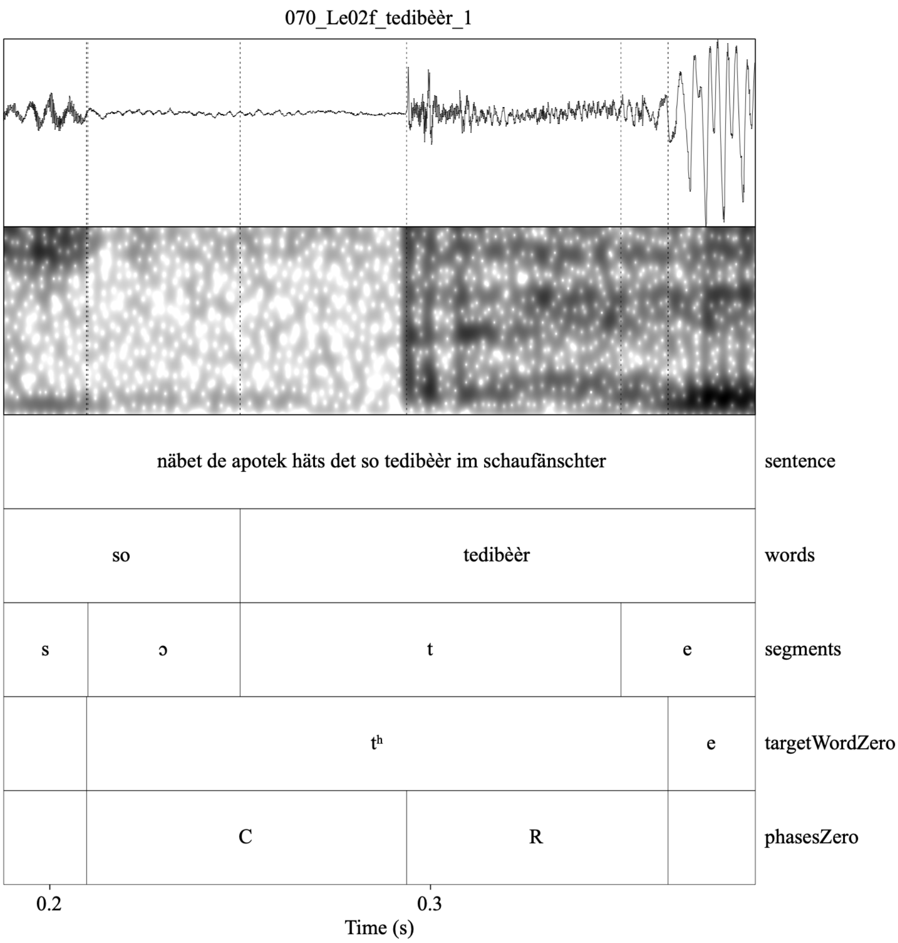


Figure 3.4: Detailed annotation of target plosive

The set criteria apply to vowels as well as any sonorant preceding the target plosive. The placement of the closure onset at the end of a clear periodic pattern results in an overall longer closure duration. For the comparison of the recording made in a soundattenuated booth, Ladd and Schmid [2018, 235] place the same boundary at the very end of a remaining voicing pattern in the wave form. However, as shown in Figure 3.4, in many cases in our recordings neither the waveform nor the spectrogram display a distinct period of complete silence of the closure phase. This might be due to an individual feature of the speaker and could also be explained by the echoing recording situation. A complete silence period can therefore not be an annotation criteria. In consequence, the criteria of putting the boundary at the end of a clear periodic pattern is an attempt to enforce consistency in the annotation of the corpus.

The annotation of the burst is mostly unambiguous because the silent and aspiration phase have drastically distinct characteristics and the boundary between them is therefore easily recognized. There are however exceptions. Seven fortis plosives /t/ and a single bilabial stop /p/ were realized as voiced plosives, e.g. *hätt s bi dier en*[8] *tach*[9] 'is there a roof on yours'. In these extreme cases, there is no obvious silent closure phase but a clear periodicity throughout both phases of the plosive. There were no clear-cut peaks visible in the waveform. Therefore, the spectrogram was used to determine the boundary placement of the burst as well as noise indication in form of a slightly shaky periodic waveform. In the same way, the reliance on the wave form to determine the offset of the release results in an overall longer segment as the annotation criteria for Zurich German plosives by Ladd and Schmid [2018]. Although the annotation primarily relies on the waveform, the spectrogram was considered more frequently to set this boundary, since the appearance of the second formant is a clear and reliant indication for the vowel [Deterding and Nolan, 2007, 387; Kleber, 2018, 471].

When the target plosive is not preceded by a vowel or a sonorant, the placement of the boundary was more challenging because of the lack of periodicity, such as a fricative in *s isch pink* 'it's pink'. As Machač and Skarnitzl [2009, 104] suggest, the decay of a fricative may be gradual and this residual noise of the fricative is faint compared to the full fricative noise. Therefore, a similar principle as for periodicity applies to the irregular acoustic pattern of fricatives where the boundary of the plosive onset is placed at the clear ending of the full fricative noise. The residual noise of the fricative is thus part of the annotated target plosive.

Plosives at the beginning of an utterance pose the biggest challenge, since it is obviously impossible to differentiate silence from the start of the closure phase of a voiceless plosive. The boundaries of the initial corpus annotation were left unchanged and only altered if they were deemed unrealistically long. Because the focus of this study relies on VOT measurements, these plosives were kept in the corpus despite the imprecise closure durations.

[8] There is a deviant gender in *en tach* 'a roof', an utterance spoken by speaker *Bu18* from the *multi* group which is a characteristic of multiethnolect Zurich German [Bruno, 2019]. The noun is not masculine but neuter which should be reflected in the indefinite article *es tach*.

[9] In this specific example, it is possible that the multiethnolect speaker intends to produce a lenis instead of a fortis plosives due to lexical interference of Standard German *das Dach* [Morand et al., 2021]. Moreover, native Zurich German speakers have shown to produce voiced lenis plosives after a nasal consonant [Ladd and Schmid, 2018, 243] which may be a contributing factor to the voicing.

In summary, the annotation of both plosive phases tends to be larger compared to other annotation standards for plosives, for example by Ladd and Schmid [2018]. The annotation is primarily based on visual cues in the sound wave and spectrogram, and verified by listening to the individual segments.

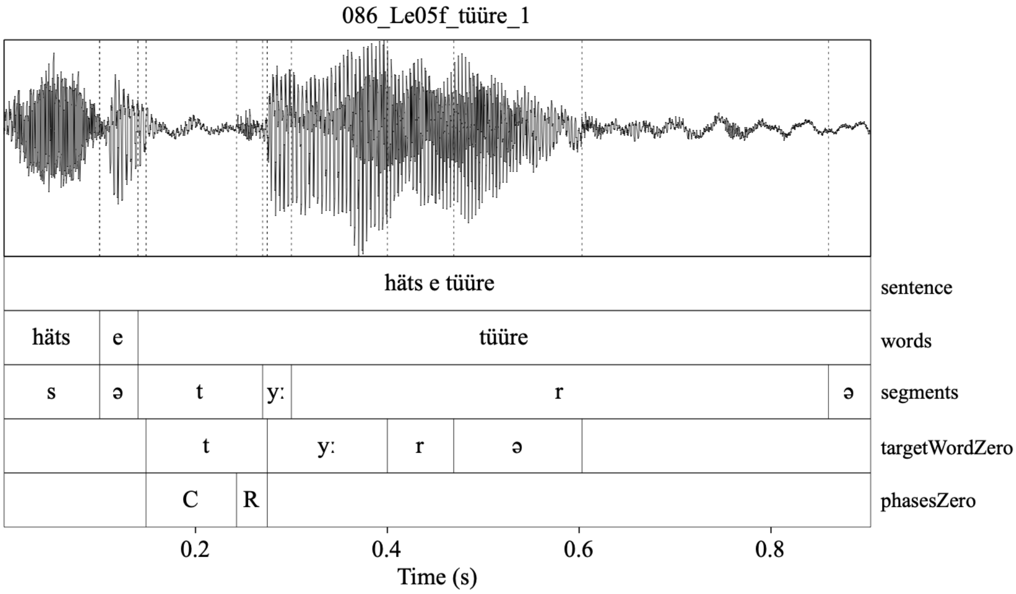


Figure 3.5: Final TextGrid file with adjusted boundaries

3.6 Durational measurements

Before taking all measurements, the Praat script in Section 9.7.1 collects the timestamps of the onset of the closure as well as the offset of the release and aligns them with the onset and offset boundaries in the *targetWord* tier. The script additionally sets all boundaries to the nearest zero crossing in the amplitude line in the tiers containing the target word as well as the tier containing the detailed plosive annotation. In this way, all annotated intervals gain more consistency since they start and end at an amplitude of zero in the waveform. As shown in Figure 3.5, the tiers are renamed to *targetWordZero* and *phasesZero*.

The most important measurement is the voice onset time of each plosive. It is calculated in the Praat script in Section 9.7.2 by subtracting the timestamp of the burst from the timestamp of the plosive offset[10]. The script also converts the unit of the calculations from seconds to milliseconds and rounds them to an integer number to facilitate the comparison and enhance legibility.

[10]The script also measures the duration of the closure phase which can be determined by the timestamp of the burst minus the timestamp of the onset of the plosive. The complete duration of each plosive is calculated by adding the durations of the closure and release of the corresponding plosive together. Moreover, the script calculates the following intensity measures in decibel: Mean intensity of the closure phase, mean intensity of release phase, mean intensity of complete plosive duration, as well as the maximum intensity in the release phase which reflects the maximum intensity of the burst. These measurements are however not further discussed in the study.

The speech rate naturally varies significantly, because the corpus consists of spontaneous speech. The script therefore additionally calculates the measure *pVOT* for articulation rate normalization. The ratio is calculated by dividing the duration of the release phase by the duration of all remaining segments of the target word following the stop consonant [Kleber, 2018, 472].

Section 9.8 in the appendix collects smaller scripts that have been written for the analysis but are not directly used for the corpus processing or plosive selection. For example, a Praat script saves all selected files from the Praat object window into a specific folder in Section 9.8.5, a Python script in Section 9.8.1 counts how many times each plosive occurs und outputs a table with all plosives sorted by the highest occurrence, or other scripts that supported renaming and moving files in Sections 9.8.2, 9.8.3, and 9.8.4. All scripts were tested on a small test corpus of 54 plosives taken from 3 speakers of each group (see Table 7.5 in the appendix).

3.7 Statistical analysis

Before discussing the method of statistical analysis, the following list shows the factors that will be used to examine the voice onset time of stop consonants in Chapter 4. Each factor in a bullet point will be compared between the three groups *mono*, *multi*, and *o60*.

- *Group:* A general comparison of all plosives in the corpus regardless of place of articulation or aspiration ([± Aspiration]) between the three groups.

- *Place of articulation (PoA):* Comparison of the stop consonants /p/ and /t/ separately between all groups regardless of their categorization in terms of [± Aspiration].

- *Aspiration:* First, a comparison of each aspirated and unaspirated plosive for each place of articulation ([p^h], [p], [t^h], and [t]) between groups.

  Additionally, a comparison between aspirated and unaspirated plosives within each group is made to test if speakers within a group produce aspirated and unaspirated plosives differently, e.g. [p^h] compared to [p] in the *mono* group.

- *Gender:* Firstly, a comparison between female and male speakers for /p/ and /t/ between each group regardless of aspiration.

  Similarly to the comparison of aspiration, there is an additional comparison between gender within each group, e.g. the difference in VOT between female and male speakers of bilabial stops in the *mono* group.

- *pVOT:* Comparison of absolute VOT and speech rate-normalized VOT (pVOT) for aspirated bilabial stops ([p^h]) and unaspirated alveolar plosives ([t]) between all groups.

- *Speaker variability:* Comparison of all three groups in regard of speaker variability within each group.

The statistical analysis was conducted in R [R Core Development Team, 2016] with the scripts in Sections 9.9.1 and 9.9.2 in the appendix. The corpus displays a lot of variability, because the data was not systematically collected for this specific study. Rather, every fortis plosive was sampled from spontaneous speech recordings. Additionally, the sample sizes per group vary considerably in all comparisons. Because of variability, the mean, standard deviation, median, and median absolute deviation are reported for each group in each comparison. The median and median absolute deviation are included because they represent a robust measure for non-normally distributed data sets. The density and quantile-quantile graphs in Figure 4.2 and Section 8.1 give an additional visual representation of each compared group. The plosive count of each group as well as all mentioned measurements are reported in Table 4.2.

The entire statistical analysis is conducted with absolute VOT measurements in milliseconds, except for the pVOT ratio analysis.

The line-up of statistical analysis was performed by implementing the outline by Levshina [2015, 171-181] for a one-way ANOVA analysis for all groups. If the ANOVA test results are significant, there are at least two groups that are different from each other per comparison. An additional post hoc test using Tukey Honest Significant Differences (HSD) gives insight to which of the *mono*, *multi*, and *o60* groups in a comparison are significantly different from one another.

Table 3.3 shows the line-up of tests that are used based on different conditions. Each comparison is conducted in four steps: The first two steps establish the conditions of the data sets of the *mono*, *multi*, and *o60* groups per comparison in terms of distribution and variance in preparation for the statistical analysis. Because of the imbalance of distribution and variance in the compared groups, different tests are applied in the analysis of the third step. The forth step conducts a post hoc test using either a parametric or non-parametric method implementing HSD.

<table>
<tr><th colspan="2">1. Normal distribution</th><th colspan="2">2. Homogeneity of variance</th><th>3. ANOVA: Test for differences between groups</th><th>4. Post hoc: Test which groups are different</th></tr>
<tr><td rowspan="4">Shapiro Wilk</td><td rowspan="2">yes</td><td rowspan="2">Bartlett</td><td>yes</td><td>Parametric one-way ANOVA</td><td>Tukey multiple comparisons of means</td></tr>
<tr><td>no</td><td>One-way analysis of means</td><td rowspan="3">Non-parametric multiple comparison test (type = Tukey)</td></tr>
<tr><td rowspan="2">no</td><td rowspan="2">Fligner-Killeen</td><td>yes</td><td>Kruskal-Wallis rank sum test</td></tr>
<tr><td>no</td><td>Asymptotic k-sample Fisher-Pitman permutation test</td></tr>
</table>

Table 3.3: Sequence of statistical tests

First, the normality of the sampling distribution is tested with the Shapiro-Wilk test[11]. The null hypothesis assumes that the sample distribution is normal and if the test turns out to be significant (p-value < 0.05) the distribution is non-normal. Non-normality is noted with *no* in Table 3.3 and respectively with *yes* for normal distribution.

The outcome of the Shapiro test determines which test for homogeneity of variance

[11] R function `shapiro.test()` in base R

is applied. The Bartlett test[12] is implemented only if all three data sets per group in a comparison are normally distributed. In any other case, the Fligner-Killeen[13] test of homogeneity of variances is used. It is a robust measure for test samples whose distribution depart from normality. Similar to the Shapiro test, the null hypothesis of the Bartlett and Fligner-Killeen tests state that the variance is homogeneous. If the test is significant (p-value smaller than 0.05) the variance across groups is not homogeneous. Heterogeneous variance is denoted as *no* in Table 3.3 and homogeneous variance as *yes*. The results of these tests are crucial to determine which test is applied to verify if the VOT of three groups are significantly different from each other.

The results of step one and two of the statistical analysis are reported in Table 4.2. All p-values are rounded to three decimal places. Some of the p-values are extremely small and are noted with < 0.001 for readability if they are in fact smaller than 0.001.

The third test per comparison (ANOVA) checks if there is a significant difference between groups. The applied test per comparison and results are listed in Tables 4.3 and 4.6. The fourth and last post hoc test shows which groups are significantly different from each other (see Tables 4.4 and 4.7 for results). The significance level of all p-values is 0.05.

If both the tests for normality of distribution and homogeneity of variance are non-significant, the parametric one-way ANOVA[14] and post hoc Tukey multiple comparisons of means[15] test is applied to determine which of the groups are significantly different from each other.

The one-way analysis of means[16] was applied to normally distributed data where the variance is not equal in all groups. The Kruskal-Wallis rank sum test[17] was used for the comparison of groups with non-normally distributions and homogeneous variances. For groups with non-normally distributions and heterogeneous variances the non-parametric ANOVA Asymptotic K-Sample Fisher-Pitman Permutation Test (R function `oneway_test()`) from the `coin` package [Hothorn et al., 2019] was used for evaluation. The non-parametric multiple comparisons (R function `nparcomp()`) of the `nparcomp` package [Konietschke et al., 2015] by implementing the HSD test (`type = "Tukey"`) was used as the post hoc test for the three mentioned tests.

The previously discussed tests are used for the comparisons of the three groups *mono*, *multi*, and *o60*. The comparisons of [± Aspiration] and gender within each group was conducted with the non-parametric Wilcoxon rank sum test that compares two data sets that do not need the be normally distributed [Levshina, 2015, 108-110]. The two-tailed test shows if two compared groups are significantly different from each other with a significance level of 0.05. The results of the Wilcoxon tests are reported in Table 4.5.

[12] R function `bartlett.test()` in base R

[13] R function `fligner.test()` in base R

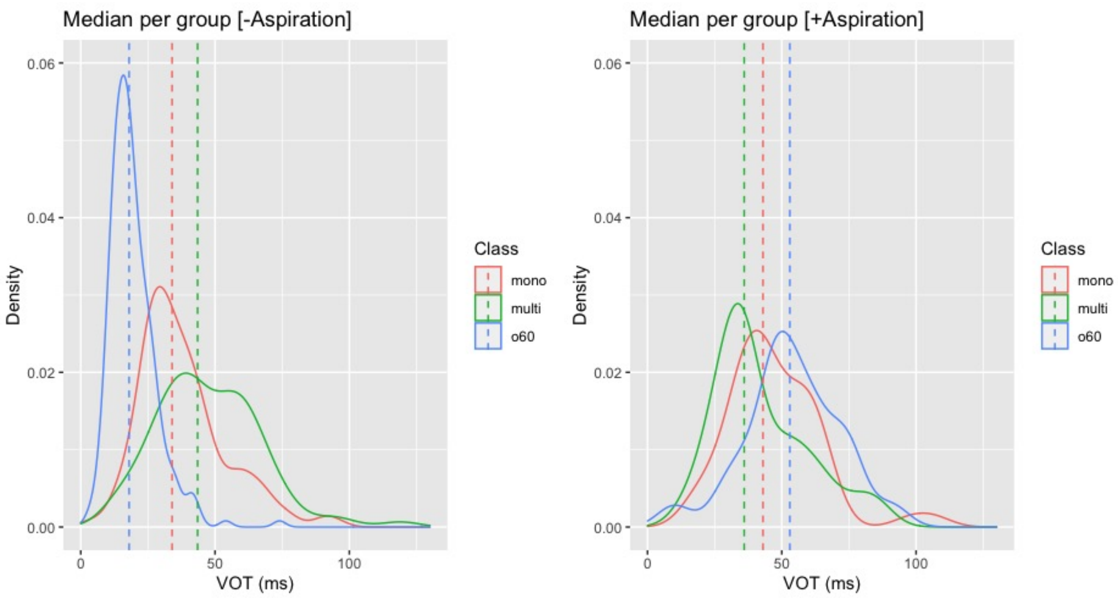
[14] R function `aov()` in base R

[15] R function `TukeyHSD()` in base R

[16] R function `oneway.test()` in base R

[17] R function `kruskal.test()` in base R

4 Results



(a) Distribution of all unaspirated stops (b) Distribution of all aspirated stops

Figure 4.1: Comparison of VOT of all plosives in terms of [±Aspiration] per group

Figure 4.1 gives a visual representation of the VOT durations per group for aspirated and unaspirated fortis plosives in the corpus separately and additionally depicts the median VOT per group as a dotted line. The median and MAD values are reported in Table 4.1. Figure 4.1a portrays the distributions of all unaspirated stops. The narrow and tall distribution shape around 15 ms of the *o60* group suggests that older Zurich German speakers produce shorter aspiration relatively consistently due to the low variability (MAD of 7.41) in the voice onset time measurements.

The distribution curves of the two younger groups are distinctly lower and broader compared to the distribution shape of the *o60* group which indicates that younger speakers produce distinctly different voice onset time in bilabial and alveolar stops compared to older speakers. The *mono* group has a considerably higher distribution paired with a lower median VOT of 34 ms in comparison to the *multi* speakers who display a median of 43.5 ms. This centered position of the *mono* group indicates that the speakers still retain some similarity in VOT production to the speakers of the *60* group but simultaneously show more similar patterns to the speakers of the *multi* group.

The *multi* speakers additionally display the highest variability score of 19.27 which positions the ethnolect speakers as most different from older Zurich German speakers. The distributions of all groups are right-skewed which hints to few outliers in every group with rather long VOT durations.

Regarding all aspirated stops in Figure 4.1b, the distribution shapes and the height of the peaks of all groups are relatively similar which is reflected by the close MAD scores (see Table 4.1).

The positions of the *o60* and *multi* groups are however reversed compared to the unaspirated stops. Whereas the *o60* group shows the highest median VOT of 53 ms, the median VOT of the multiethnolect speakers is the shortest with a median VOT of 36 ms. The *mono* group is again positioned between the other two groups with a median VOT of 43 ms. The medians of both younger groups are closer together and distinctly distant from the median of the *o60* group in both graphs.

	Aspiration	Plosive count	VOT median	VOT mad
mono	+	53	43	14.83
	−	110	34	12.6
multi	+	34	36	12.6
	−	86	43.5	19.27
o60	+	23	53	14.83
	−	220	18	7.41

Table 4.1: VOT measurements and token count per group

4.1 Results of VOT measurements

Table 4.2 displays the token count of each individual group, as well as the VOT durations in mean, standard deviation (SD), median, and median absolute deviation (MAD) of all group comparisons. These are titled as *group*, *place of articulation (PoA)*, *aspiration*, and *gender* in Table 4.2. The figures report absolute measurements of VOT in milliseconds that are non-normalized. Speech rate normalization is discussed in Section 5.4. The table additionally includes the results of the tests for normal distribution and homogeneity of variance.

Table 4.3 lists the results of the non-parametric ANOVA and alternative tests that determine if there is a significant difference between groups. Lastly, the findings of the post hoc tests per comparison determine which groups have significantly different VOT. These are reported in Table 4.4. The most important result of every test for this analysis is the p-value. The significance level of all tests is 0.05. In consequence, any p-value below 0.05 represents a significant result.

Factor		Group	Plosive count	VOT (ms) Mean	SD	Median	MAD	p-value Shapiro	Bartlett	Fligner	Normal Distribution	Homogeneity of Variance
Group		mono	163	40.85	16.98	38	14.83	< 0.001		< 0.001	no	no
		multi	120	46.03	18.83	42	17.79	0.001				
		o60	243	22.72	14.14	18	7.41	< 0.001				
		total	526									
Place of articulation	/p/	mono	48	46.54	18.18	42.5	14.83	0.006		0.558	no	yes
		multi	38	41.26	16.73	38.5	15.57	0.194				
		o60	33	31.18	22.36	26	16.31	0.002				
		total	119									
	/t/	mono	115	38.48	15.94	35	13.34	< 0.001		< 0.001	no	no
		multi	82	48.23	19.42	44.5	20.76	0.007				
		o60	210	21.39	11.91	18	7.41	< 0.001				
		total	407									
Aspiration	[p^h]	mono	41	45.73	18.31	42	13.34	0.002		0.465	no	yes
		multi	25	43.36	17.29	37	14.83	0.035				
		o60	13	50.85	22.75	48	22.24	0.82				
		total	79									
	[p]	mono	7	51.29	17.99	51	22.24	0.97	0.039		yes	no
		multi	13	37.23	15.44	42	14.83	0.325				
		o60	20	18.4	8.85	17	9.64	0.563				
		total	40									
	[t^h]	mono	12	51.5	11.09	52	14.08	0.504	0.044		yes	no
		multi	9	39.33	18.96	34	4.45	0.056				
		o60	10	57.5	8.05	58	7.41	0.681				
		total	31									
	[t]	mono	103	36.96	15.76	33	11.86	< 0.001		< 0.001	no	no
		multi	73	49.33	19.32	47	20.76	0.014				
		o60	200	19.59	8.78	18	7.41	< 0.001				
		total	376									
Gender	/p/ f	mono	33	50.48	18.9	44	16.31	0.005		0.244	no	yes
		multi	19	44.37	14.09	40	10.38	0.071				
		o60	24	38	22.34	34	21.5	0.029				
		total	76									
	/p/ m	mono	15	37.87	13.28	37	17.79	0.359	0.021		yes	no
		multi	19	38.16	18.88	37	20.76	0.38				
		o60	9	13	7.28	13	5.93	0.883				
		total	43									
	/t/ f	mono	86	38.05	16.83	33	11.86	< 0.001		< 0.001	no	no
		multi	24	55.46	25.36	53.5	25.95	0.394				
		o60	120	20.93	12.34	17.5	6.67	< 0.001				
		total	230									
	/t/ m	mono	29	39.76	13.13	41	8.9	0.41		< 0.001	no	no
		multi	58	45.24	15.67	42.5	18.53	0.191				
		o60	90	22.01	11.34	20	7.41	< 0.001				
		total	177									

Table 4.2: Plosive count, VOT measurements, and tests for normal distribution and homogeneity of variance per compared factor.

		One-way analysis of means				Kruskal-Wallis rank sum test			Asymptotic K-Sample Fisher-Pitman Permutation Test		
Factor		F	df	denom df	p-value	Chi-squared	df	p-value	Chi-squared	df	p-value
	Group								151.88	2	< 0.001
PoA	/p/					15.268	2	< 0.001			
	/t/								145.9	2	< 0.001
Aspiration	[p^h]					1.447	2	0.485			
	[p]	15.983	2	12.91	< 0.001						
	[t^h]	3.781	2	16.325	0.045						
	[t]								166.48	2	< 0.001
Gender	/p/ f					7.144	2	0.028			
	/p/ m	23.204	2	26.498	< 0.001						
	/t/ f								80.355	2	< 0.001
	/t/ m								71.776	2	< 0.001

Table 4.3: Results of non-parametric ANOVA show if there is a significant difference in VOT between groups.

			No-parametric multiple comparison test				
Factor		Groups	Estimator	Lower	Upper	Statistic	p-value
Group		mono, multi	0.584	0.501	0.662	2.387	0.046
		mono, o60	0.16	0.118	0.214	-11.067	0
		multi, o60	0.132	0.092	0.186	-10.835	0
PoA	/p/	mono, multi	0.415	0.28	0.564	-1.332	0.373
		mono, o60	0.257	0.14	0.424	-3.276	0.002
		multi, o60	0.307	0.176	0.478	-2.602	0.024
	/t/	mono, multi	0.654	0.555	0.741	3.621	< 0.001
		mono, o60	0.153	0.107	0.214	-9.841	0
		multi, o60	0.095	0.058	0.151	-10.122	0
Aspiration	$[p^h]$	mono, multi	0.448	0.285	0.623	-0.686	0.774
		mono, o60	0.582	0.346	0.785	0.796	0.706
		multi, o60	0.603	0.359	0.805	0.975	0.591
	[p]	mono, multi	0.269	0.071	0.639	-1.502	0.316
		mono, o60	0.043	0.003	0.379	-2.806	0.014
		multi, o60	0.171	0.048	0.46	-2.624	0.024
	$[t^h]$	mono, multi	0.25	0.057	0.65	-1.469	0.29
		mono, o60	0.65	0.349	0.865	1.145	0.476
		multi, o60	0.789	0.351	0.963	1.565	0.246
	[t]	mono, multi	0.7	0.597	0.786	4.407	< 0.001
		mono, o60	0.129	0.085	0.191	-9.736	< 0.001
		multi, o60	0.058	0.03	0.11	-9.536	< 0.001
Gender	/p/ f	mono, multi	0.404	0.235	0.598	-1.151	0.478
		mono, o60	0.306	0.159	0.506	-2.258	0.059
		multi, o60	0.348	0.18	0.563	-1.655	0.216
	/p/ m	mono, multi	0.479	0.259	0.707	-0.206	0.994
		mono, o60	0.044	0.006	0.267	-3.519	0.001
		multi, o60	0.082	0.018	0.299	-3.646	< 0.001
	/t/ f	mono, multi	0.724	0.559	0.845	3.103	0.557
		mono, o60	0.151	0.097	0.228	-8.027	< 0.001
		multi, o60	0.082	0.033	0.188	-5.938	< 0.001
	/t/ m	mono, multi	0.579	0.427	0.717	1.23	0.477
		mono, o60	0.146	0.068	0.287	-4.867	< 0.001
		multi, o60	0.101	0.053	0.184	-7.423	< 0.001

Table 4.4: Results of post hoc test identify groups that are significantly different from one another.

			Wilcoxon rank sum test				
				95% confidence interval		difference	
	Group	Contrast	W	lower	upper	in location	p-value
Aspiration	mono	[p], $[p^h]$	109	-20.999	8.999	-7.999	0.32
		[t], $[t^h]$	991.5	8	24	16	<0.001
	multi	[p], $[p^h]$	180	-7	17.999	4.999	0.59
		[t], $[t^h]$	228.5	-24.999	3	-10	0.138
	o60	[p], $[p^h]$	239	19	45.999	30	<0.001
		[t], $[t^h]$	1986	33	43.999	38.999	<0.001
Gender	mono	/p/ : f, m	338	< 0.001	20.999	9.999	0.044
		/t/ : f, m	1031.5	-10	2	-4.442	0.165
	multi	/p/ : f, m	223.5	-4	17.999	7	0.209
		/t/ : f, m	861	-1	18.999	7.217	0.092
	o60	/p/ : f, m	192.5	9	34	20	<0.001
		/t/ : f, m	4750	3.999	0.999	-1.999	0.136

Table 4.5: Results of pairwise comparison tests if VOT is produced significantly different in aspirated and unaspirated stops within groups.

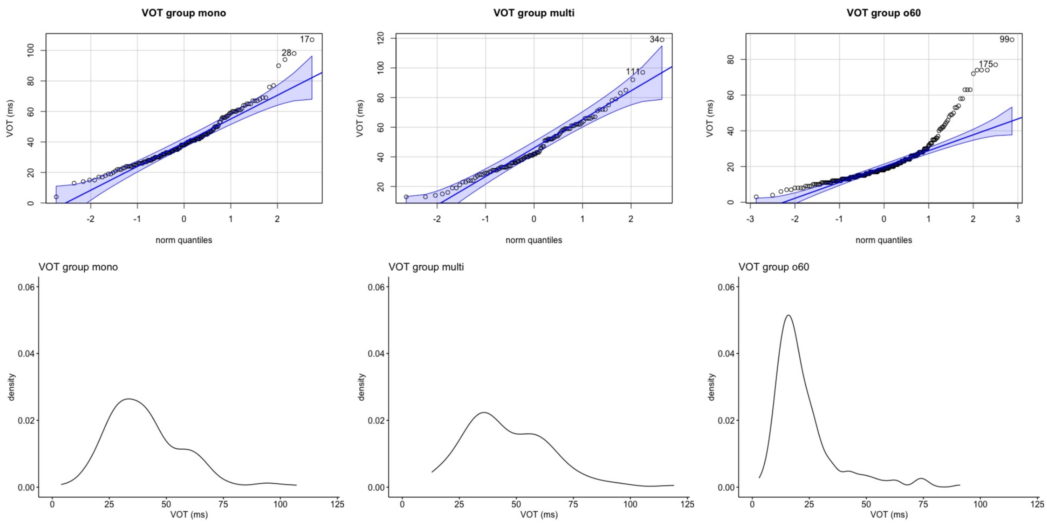


Figure 4.2: Quantile-quantile and density plots for VOT in the *mono*, *multi*, and *o60* groups

4.2 Group comparison

This section discusses all plosives in the corpus collectively per group in more detail regardless of the categorization in terms of aspiration or place of articulation. The top half of Figure 4.2 shows the quantile-quantile plots of all fortis plosives in the corpus separated into the three groups *mono*, *multi*, and *o60*. The lower half displays the corresponding distribution plots. The quantile-quantile plots outline a correlation line between the data of the corpus and a normal distribution where each data point, i.e. each VOT measurement, is assigned a quantile and mapped in the plot with a corresponding quantile of a normal distribution. The plots therefore give insight, supplementary to the density plots, if the VOT measurements per group are normally distributed. Similarly to the right-skewed density plots, all quantile-quantile plots confirm that VOT measurements in each group are non-normally distributed since the data points in each plot are not strictly positioned along the correlation line, and diverge especially in higher VOT durations.

The quantile-quantile plot of older Zurich German speakers on the right shows the most outliers starting around a VOT of 30 ms that strongly depart from an expected normal distribution which can also be observed in the long tail of the right-skewed density plot. This suggests that the *o60* speakers produce, on the one hand, relatively long VOT durations and on the other, relatively similar shorter VOT durations. This could indicate that the *o60* group produces stop consonants distinctly in terms of [±Aspiration] as previously reported in Figure 4.1.

The VOT measurements of the *mono* speakers in the quantile-quantile plot on the left show far less outliers in longer VOT measurements than the other two groups. This might point to a similar distinction between aspirated and unaspirated stops of the *o60* group, but there is notably less indication of a distinction in terms of [±Aspiration] compared to older speakers.

The *multi* group seems to have the least outliers since only few data points are far off the correlation line. In consequence, there is no indication that multiethnolect speakers clearly differentiate between aspirated and unaspirated stop consonants based on a first visual inspection of the VOT measurements.

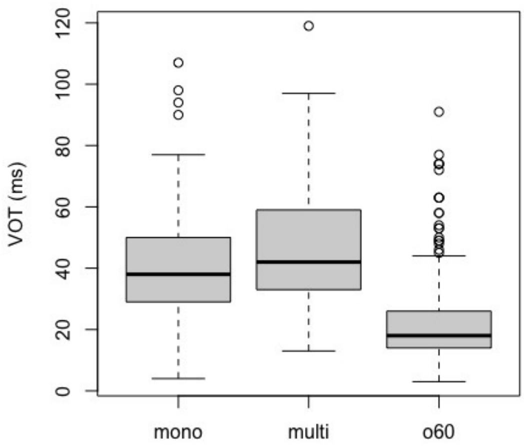


Figure 4.3: VOT comparison of all plosives per group

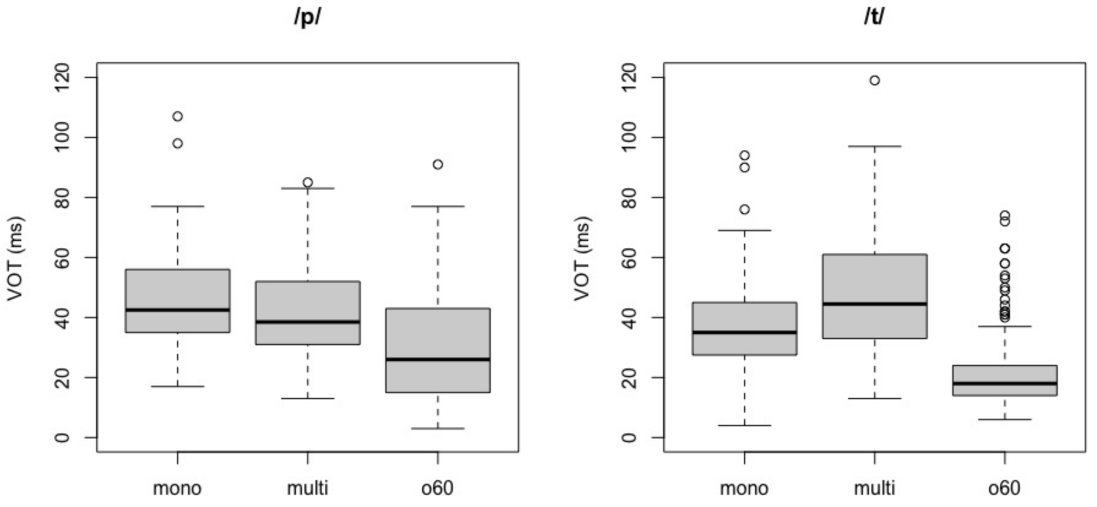
Figure 4.3 visualizes the measured VOT per group as boxplots based on quantiles and the median. As listed in Table 4.2, the *multi* group has notably the highest VOT mean with roughly 46 ms[1] and median with 42 ms of all groups. However, the data set containing only 120 plosives is also the smallest. The group displays also the highest standard deviation (nearly 19) and MAD (nearly 18).

The voice onset time durations of older Zurich speakers, on the other hand, measure less than half of the multiethnolect group with a mean of nearly 23 ms and a median of 18 ms. The *o60* group also shows a considerable lower SD (14) and MAD (7). The *o60* group counts around double the amount of plosives compared to the multiethnolect group.

All VOT measurements of the *mono* group are situated between the other two groups but they are closer to the *multi* group with a mean of almost 41 ms and median of 38 ms. This fact is clearly illustrated in Figure 4.3 where the boxplots of both groups of younger speakers are more similar and clearly distinct from the boxplot of older speakers.

The variance across the groups is heterogeneous. The p-value of the non-parametric ANOVA test is smaller than 0.001 which confirms that there is a significant difference in the release durations of plosives across the *mono*, *multi*, and *o60* groups (see Table 4.3). The post hoc test determines a considerable difference between the *o60* and *mono* groups as well as between the *o60* and *multi* groups as seen in Table 4.4. The difference between the *mono* and *multi* groups is significant with a p-value of 0.046 just under the significance level of 0.05.

[1] The mean measurements in Table 4.2 are rounded to two decimal digits, but rounded to integers in the text for readability.



(a) VOT of all bilabial plosives per group (b) VOT of all alveolar plosives per group

Figure 4.4: Comparison of VOT durations of bilabial and alveolar plosives

4.3 Place of articulation

In this section, the results of VOT according to place of articulation of /p/ and /t/ are reported regardless of their categorization in terms of aspiration. As displayed in Figure 4.4a[2], both groups with younger Zurich German speakers produce considerably higher VOT durations than the older speakers.

However, the *mono* speakers have the highest mean and median (almost 47 ms and 43 ms) for the stop consonant /p/ compared to the other two groups as reported in Table 4.2. With the biggest plosive count of 48 bilabial stops, they also display the least variability of all three groups with a MAD value of nearly 15. The older speakers show again the smallest mean and median with 31 ms and 26 ms release duration. This group also has the smallest sample size of 33 bilabial stops and the highest amount of variability with a standard deviation of 22 and MAD of 16. The VOT measurements of bilabial stops by the *multi* speakers are set between the other groups but closer located to the younger speakers of the *mono* group with a mean of 41 ms and median of 38.5 ms, and a standard deviation of nearly 17 and a median absolute deviation of nearly 16.

Because the distribution across all groups is non-normal and variance is homogeneous, the Kruskal-Wallis test was applied and turned significant. The p-value of 0.373 of the post hoc test determines that there is no significant difference between the *mono* and *multi* groups. However, the VOT durations of the older speakers are significantly different from

[2]The density and quantile-quantile plots of each group are displayed in Section 8.1 in the appendix for each comparison, but won't be discussed further.

both younger groups.

The differences between all groups for the alveolar stop consonant /t/ are shown in 4.4b. The multiethnolect speakers display the longest VOT in Table 4.2 with a mean VOT of 48 ms and median of 44.5 ms. This group also has the highest standard deviation and MAD (both around 20) and therefore the biggest variability in production of VOT. However, they have by far the smallest token count of 82 alveolar stops compared to 115 stops in the *mono* group. The *o60* group count over two and half as many alveolar stops with 210 tokens.

The *o60* speakers have the shortest mean (21 ms) and median (18 ms) voice onset time, as well as the smallest variability with a standard deviation of nearly 12 and a median absolute deviation of 7.

The VOT measurements of the *mono* group regarding alveolar stops is set between the other two groups, but is closer to the multiethnolect speakers with a mean of 38 ms and median VOT duration of 35 ms.

The tests for normality in distribution and homogeneity of variance turned both significant. The resulting non-parametric ANOVA test detected a significant difference across the groups. The following post hoc test confirms that every group produces VOT significantly different from one another (see Tables 4.3 and 4.4).

The token count per group in every comparison is rather unbalanced. There are less than half as many bilabial (48) than alveolar stops (115) in the *mono* group, and about half as many bilabials (38) compared to alveolar stops (82) in the multiethnolectal group. The biggest difference is found in the *o60* group with 33 bilabial stops compared to 210 alveolar stops. Thefore, there is a notable bias towards alveolar stops in the corpus.

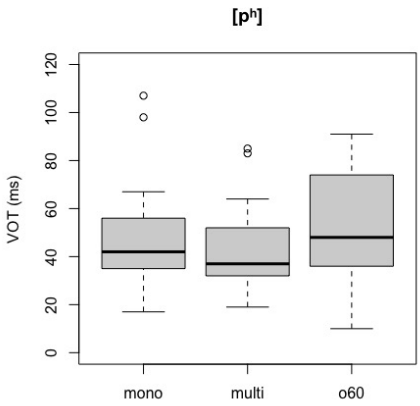
4.4 Aspiration

The following section differentiates the VOT measurements of plosives in terms of place of articulation and aspiration. As apparent in Figures 4.5a and 4.5c, the *o60* group has the highest values in aspirated plosives [p^h] and [t^h] with means and medians around 50 ms for [p^h] and 58 ms for [t^h] (see Table 4.2). These are also the longest VOT measurements of all compared samples. Additionally, those are the only two instances where the older Zurich German speakers do not have the shortest VOT of all groups per comparison.

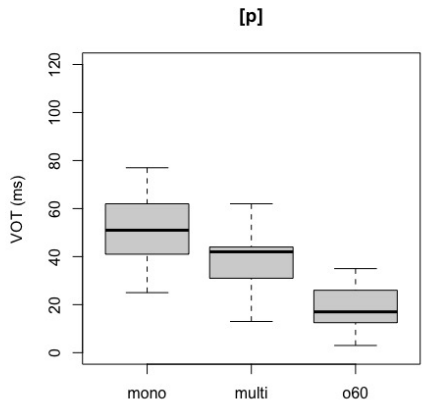
The means (43 and 39 ms) and medians (37 and 34 ms) of the *multi* group for aspirated plosives [p^h] and [t^h] are considerably shorter compared to the *o60* group. The VOT measurements of the *mono* group are yet again set between the other two groups in both comparisons of aspirated stops. The means and medians of VOT for [p^h] are 46 ms and 42 ms, and for [t^h] both around 52 ms.

Regarding the comparison of unaspirated bilabial stops [p^h], the distribution is non-normal and the variance is homogeneous. The p-value of the resulting Kruskal-Wallis test is 0.485 as reported in Table 4.3. With a significance level of 0.05, the test is clearly not significant which means that there is no difference between the production in VOT in aspirated bilabial stops. The following post hoc test is technically redundant, but the high p-values (0.774, 0.706 and 0.591 in Table 4.4) for all group comparisons nevertheless confirm that there is no significant difference between groups. In other words, all three groups produce aspirated bilabial stops with similar voice onset time measurements.

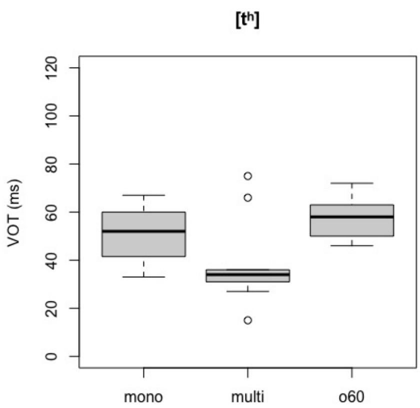
The distribution of the aspirated alveolar stop [t^h] is normal and the variance is homogeneous. The one-way analysis of means returns a barely significant p-value of 0.045 with a significance level of 0.05 (see Table 4.3). This result suggests that there is a significant difference between VOT production in aspirated alveolar stops [t^h] across the three groups. The high p-values of post hoc test (0.29, 0.476, and 0.246 in Table 4.4) however suggest that there is no significant difference in VOT production. In consequence, this indicates that all groups have comparable VOT durations in aspirated alveolar stops.



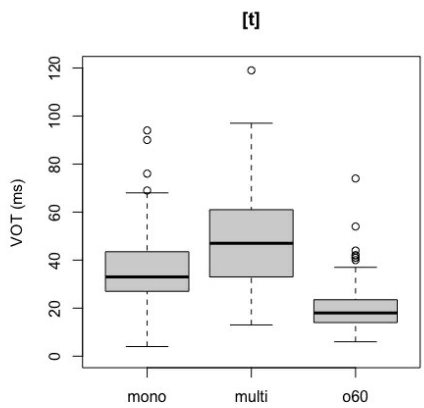
(a) Comparison of aspirated bilabial stops



(b) Comparison of unaspirated bilabial stops



(c) Comparison of aspirated alveolar stops



(d) Comparison of unaspirated alveolar stops

Figure 4.5: Group comparison of aspirated and unaspirated bilabial and alveolar stops

As shown in Figure 4.5b, the *mono* group clearly shows the longest VOT in unaspirated bilabial stops [p] with a mean and median of 51 ms (see Table 4.2). The VOT measurements of the unaspirated bilabial stops are unexpectedly higher than the measurements of the aspirated stops [p^h]. Further, the *mono* group has the highest variability with a standard deviation of nearly 18 and median absolute deviation of 22. The sample size consists of merely 7 unaspirated plosives. The other younger speakers of the *multi* group have a considerably lower mean (37 ms) and median (42 ms) compared to the *mono* group. The VOT measurements of the older speakers are distinctly shorter with a mean of 18ms and median of 17 ms compared to the younger speakers. They also display clearly lower variability with a standard deviation and median absolute deviation around 9.

The distribution across all three groups is normal and the variance is heterogeneous. The one-way analysis of means returns a significant result and the post hoc tests show that there is no difference in younger Zurich German speakers in the production of unaspirated bilabial stops (the p-value is 0.316 see Table 4.4), but the p-values of the comparison of the *mono* and *o60* groups (0.014), and the *multi* and *o60* groups (0.024) are below the significance level. Therefore, there is a significant difference between younger and older speakers in the production of unaspirated bilabial stops and unaspirated alveolar stops.

The longest VOT durations for unaspirated alveolar stops [t] is produced by the multiethnolect group (see Figure 4.5d) with a mean VOT duration of 49ms and median of 47 ms as shown in Table 4.2. The group also displays the highest variability with a SD value of 19 and MAD of nearly 21. The *o60* speakers have the shortest VOT that lasts even less than half the VOT duration of the *multi* group with a mean of 20 ms and a median of 18 ms. With a standard deviation of nearly 9 and MAD of 7, the variability in this group is also considerably lower compared to the multiethnolect speakers. The mean and median VOT measurements of the *mono* group are located practically in the middle of the other groups with a mean VOT of 37 ms and median of 33 ms.

The test for normality of distribution and homogeneity of variance, as well as the following non-parametric tests are all significant. The results of the post hoc test show that there are statistically significant differences between all groups, i. e. all groups produce significantly different VOT from one another.

The plosive count of all three groups are relatively small for the comparisons [p^h], [p], and [t^h]. The lowest count is only 7 for unaspirated bilabial stops in the *mono* group and range up to 41 aspirated bilabials in the *mono* group. By comparison, all token counts in the comparison of unaspirated alveolar stops are rather high with 103 stops in the *mono* group, 73 stop in the *multi* group, and lastly the by far biggest count of 200 stops in the *o60* group.

The above presented results are discussed in regard of group comparison. The following paragraph describes the comparison of VOT within each group in terms of place of articulation and aspiration[3]. For example, some mean and median VOT measurements

[3]The boxplots for each pairwise comparison of aspirated and unaspirated stops for each group are in Section 8.2 in the appendix.

between unaspirated and aspirated stops in both places of articulation within groups are close together, e.g. VOT for [p] (37 and 42 ms) and [p^h] (43 and 37ms) in the *multi* group. The comparison will therefore be determined by using the Wilcoxon rank sum test if the *mono* group produces significantly different aspirated and unaspirated bilabial stops, e.g. VOT measurements of [p] compared to VOT measurements of [p^h] within each group. The p-values with a significance level of 0.05 are reported in Table 4.5. The mean and median values won't be repeated at this point, and the following paragraph lists the results of the statistical analysis.

The low p-values in the two comparisons of aspirated and unaspirated stops of the *o60* group suggest that older Zurich German speakers produce aspirated and unaspirated plosives distinctly different regardless of place of articulation.

The comparison of alveolar stops in terms of aspiration in the *mono* group returns a significant p-value which indicates that *mono* speakers make a distinction in the production of unaspirated and aspirated alveolar plosives. However, the test is not significant for the VOT measurements of bilabial stops with a p-value of 0.32. The *mono* group therefore seems to make no distinction in the articulation of unaspirated and aspirated bilabial plosives.

The p-values of the Wilcoxon test for both comparisons of the *multi* group are not significant (p-values of 0.59 and 0.138). In consequence, the multiethnolect speakers seem to produce unaspirated and aspirated bilabial and alveolar stops with similar voice onset time that are not statistically distinct. In other words, multiethnolect speakers do not distinguish between aspirated and unaspirated stops.

4.5 Gender

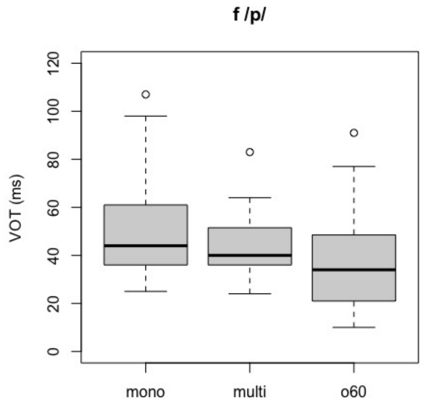
This section describes the results of the group comparisons in regard of gender and place of articulation. The comparison does not distinguish between aspiration because the token counts per group especially for unaspirated bilabials and aspirated alveolars would be considerably below 10 stops per group which are too few for a comparison.

Figure 4.6a shows that female speakers of the *mono* group have the longest VOT in bilabial plosives with a mean VOT of 50 ms and median of 44 ms compared to the other groups (see Table 4.2).

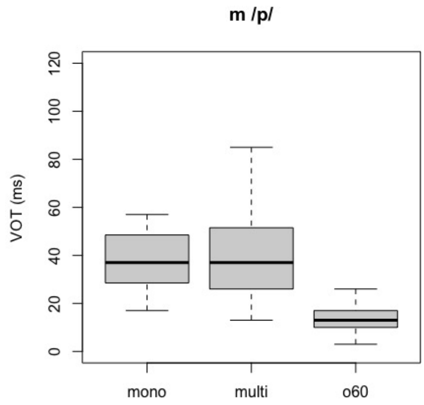
The older female Zurich German speakers produce the bilabial stop consonant /p/ with a mean VOT of 38 ms and median of 34 ms. They also show the highest variability with 22.34 standard deviation and 21.5 median absolute deviation.

The female multiethnolect speakers are placed right in the middle of the other two groups with a mean of 44 ms and a median of 40 ms. The variability however is higher with a SD of 14 and MAD of 10 compared to the other two groups of female speakers.

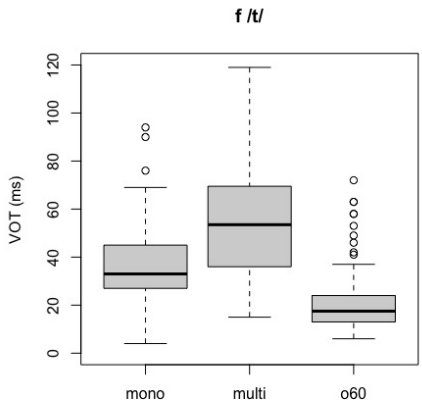
The distribution across all groups is non-normal and the variance is homogeneous. The Kruskal-Wallis test returns a significant p-value of 0.028 (see Table 4.3). None of the results of the post hoc tests are however significant. This suggests that female speakers do not produce significantly different VOT in bilabial stops. The post hoc test of the comparison between the *mono* and *o60* groups however returns a p-value of 0.059 as reported in Table 4.4. With a significance level of 0.05, the results are just barely not statistically significant.



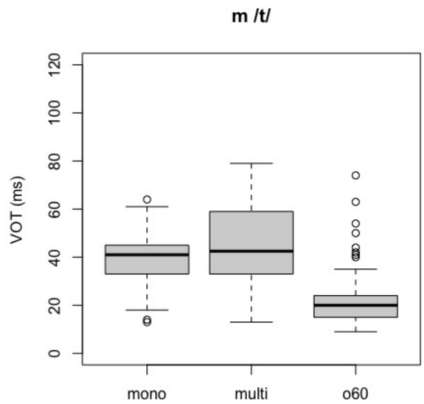
(a) Females: /p/



(b) Males: /p/



(c) Females: /t/



(d) Males: /t/

Figure 4.6: Gender differences of in VOT measurements

As apparent in Figure 4.6b, the male speakers of both adolescent groups produce the bilabial plosive very similarly. When rounded, they have the same VOT measurements with a mean of 38 ms and median of 37 ms as listed in Table 4.2. The multiethnolect speakers show considerably higher variability with a standard deviation of nearly 19 and MAD of nearly 21, whereas the SD of the *mono* is 13 and MAD almost 18.

The male *o60* speakers have notably lower mean and median values (both 13 ms) which is the shortest VOT measurements in all comparisons. The variability is also the lowest with a SD of 7 and MAD of nearly 6.

The token count in all groups are fairly low with 15 bilabial stops in the *mono* group, 19 stops in the *multi* group, and only 9 plosives in the *o60* group. The distribution is normal and the variance is heterogeneous. The results of the one-way analysis of means is significant which means there is a significant difference in VOT production by male speakers across the three compared groups. The high p-value (0.994) of the comparison between the male adolescent groups suggests that there is no durational difference in voice onset time (see Table 4.4). The other two post hoc tests however show evidence for different VOT durations between the *mono* and *o60* group, as well as the *multi* and *o60* groups. This suggests that there is a generational difference in male speakers in the articulation of bilabial stops.

Regarding the alveolar stops as shown in Figure 4.6c, the older female speakers display the shortest VOT duration with a mean of 21 ms and median of 17.5 ms (see Table 4.2). The variability is considerably lower compared to the two adolescent groups with a standard deviation of 12 and MAD of nearly 7. The *mono* female speakers produce VOT measurements of almost double the length in mean (38 ms) and median (33 ms). The voice onset time of female multiethnolects is even higher with a mean of 55 ms and median of 53.5 ms. They also have a rather high standard deviation (25 ms) and median absolute deviation (almost 26 ms).

Figure 4.6d shows that, similarly to the female speakers of the *o60* group, the older male speakers have a considerably lower mean VOT (22 ms) and median VOT (20 ms) with lower variability scores (SD of 11 ms and MAD of 7 ms) compared to the male speakers of both younger groups.

The mean (nearly 40 ms) and median (41 ms) of the male *mono* speakers are around twice the length of the older male speakers. The longest release durations are produced by the *multi* group with a mean of 45 ms and median of 42.5 ms. The male multiethnolects also have a slightly higher variability compared to the other two groups with standard deviation of nearly 16 and MAD of nearly 19.

For both female and male comparisons of alveolar stops, the distribution is non-normal and the variance is heterogeneous across the respective groups. Both comparisons return significantly low p-values in the non-parametric ANOVA test (see Table 4.3). The post hoc tests show that there are no significant differences in VOT measurements between younger speakers in the *mono* and *multi* groups in each compared gender category with a p-value of 0.557 for female speakers and a p-value of 0.477 for male adolescents. There are however significant differences between older and younger speakers in each gender category. Again, the post hoc test shows evidence for a generational difference in VOT production in alveolar stops for female and male speakers.

The results above describe the comparison between groups. The following paragraph addresses the differences within groups in regard of place of articulation and gender, i.e. if female and male speakers within a group articulate the voice onset time of a bilabial or alveolar stop differently. The durational measurements of VOT are not repeated here as they are mentioned in the paragraphs above[4].

The results of the Wilcoxon rank sum test are reported in the section titled *gender* in Table 4.5. The p-value of 0.044 of the comparison of bilabial stops of female and male

[4]The boxplots for each comparison are in Section 8.3 in the appendix.

speakers in the *mono* group is barely significant with a significance level of 0.05. This suggests that female and male speakers in the *mono* group produce bilabial stops with significantly distinct VOT. The Wilcoxon test shows however no significant result for the same group regarding the alveolar stop with a p-value of 0.165.

The results for the *multi* group are not significant with p-values of 0.209 for bilabial stops and 0.092 for alveolar stops. The test therefore does not provide evidence that female and male multiethnolect speakers produce distinctly different VOT for the two examined plosives.

The low p-value (< 0.001) for the comparison of bilabial stops between female and male speakers of the *o60* group is significant. The result of the alveolar stop however is not significant with a p-value of 0.136. In consequence, female and male speakers of the *o60* group seem to produce distinct VOT in bilabial stops but not in alveolar stops.

4.6 Speech rate normalization

The composition of this corpus does not allow to compare the individual target words since there is not enough data. Most target words are not even produced by every speaker. Instead, the pVOT comparisons were conducted on target words with an equal syllable count to ensure that the compared words are of similar length. Target words with more than two syllables were excluded due to an increased number of phonetic units per word and scarcity in corpus. In order to make the compared words as similar as possible, the comparison differentiates between place of articulation of stop consonants as well as between aspirated and unaspirated stops. Unfortunately, there was not enough data to compare monosyllabic unaspirated bilabial stops, as there are only 9 tokens in all three groups combined. The same applies to monosyllabic words with aspirated alveolar stops which amount to a combined count of merely 6 stops.

Table 4.6 shows the plosive count, VOT calculations, as well as the results of the test for normality of distribution and homogeneity of variance for both the absolute measurements of VOT and the normalized pVOT for the comparisons of aspirated bilabial stops [p^h] and unaspirated alveolar stops [t].

Regarding monosyllabic target words with a word-initial aspirated bilabial stops [p^h] in the corpus, the *o60* group show a rather high mean VOT of 69.5 ms and median VOT of 74 ms compared to both groups of younger speakers. The variability is very low with a mean of 11 and MAD of 2 with a small plosive count of 4. The *multi* and *mono* groups both have a plosive count of 11 and median VOT of 47 ms. Whereas the MAD for the *mono* group is nearly 15, the variability for the multiethnolect speakers is even higher with a MAD VOT of 22.

In the speech rate normalized calculations for the aspirated bilabial stop, the *multi* group show the lowest median pVOT of 0.14 and the *mono* group the highest median pVOT of 0.18. The *o60* group is situated precisely in the middle with a median pVOT of 0.16. The variability MAD measures for pVOT for the *o60* and the *mono* group are rather low with 0.045 and 0.03 compared to the median absolute deviation pVOT 0.11 of multiethnolect speakers.

As shown in Table 4.7, the statistical comparison using the parametric one-way ANOVA for the comparison of pVOT and the Kruskal-Wallis rank sum test for the comparison of VOT measurements are both non-significant. These results already demonstrate that there are no statistically significant differences between the three groups. The p-values of the post hoc test are all clearly above the significance level of 0.05 and therefore confirm that all three speaker groups produce VOT distinctly different from one another in aspirated bilabial stops regardless of speech rate in monosyllabic target words.

Factor		Groups	Plosive count	VOT				p-value			Normal Distribution	Homogeneity of Variance
				Mean	SD	Median	MAD	Shapiro	Bartlett	Fligner		
[pʰ]	pVOT	mono	11	0.17	0.04	0.18	0.045	0.683				
		multi	11	0.15	0.07	0.14	0.11	0.267	0.207		yes	yes
		o60	4	0.16	0.03	0.16	0.03	0.491				
	VOT	mono		54.27	18.14	47	14.83	0.049				
		multi		50.45	21.75	47	22.24	0.53		0.28	no	yes
		o60		69.5	11.09	74	2.22	0.03				
[t]	pVOT	mono	68	0.19	0.09	0.17	0.08	0.008				
		multi	48	0.24	0.12	0.23	0.12	0.054		< 0.001	no	no
		o60	86	0.08	0.05	0.06	0.04	< 0.001				
	VOT	mono		38.25	16.97	35	13.34	< 0.001				
		multi		49.19	20.5	52	20.76	0.027		< 0.001	no	no
		o60		19.47	10.28	17	5.93	< 0.001				

Table 4.6: Comparison of absolute (VOT) and normalized (pVOT) voice onset time

Factor		Groups	ANOVA			Post hoc				
[pʰ]	pVOT		parametric one-way ANOVA			Tukey multiple comparisons of means				
			df	F value	Pr(>F)	diff	lower	upper	p adj	
		mono, multi	2	0.442	0.648	-0.021	-0.078	0.035	0.621	
		mono, o60				-0.012	-0.089	0.066	0.926	
		multi, o60				0.01	-0.068	0.087	0.948	
	VOT		Kruskal-Wallis rank sum test			non-parametric comparison				
			Chi-squared	df	p-value	estimator	lower	upper	statistic	p-value
		mono, multi	3.338	2	0.188	0.426	0.172	0.725	-0.556	0.887
		mono, o60				0.818	0.373	0.971	1.751	0.195
		multi, o60				0.75	0.355	0.942	1.523	0.294
[t]	pVOT		non-parametric ANOVA			non-parametric comparison				
			Chi-squared	df	p-value	estimator	lower	upper	statistic	p-value
		mono, multi	80.144	2	< 0.001	0.631	0.494	0.749	2.252	0.064
		mono, o60				0.108	0.06	0.186	-7.842	< 0.001
		multi, o60				0.095	0.047	0.184	-6.971	< 0.001
	VOT		non-parametric ANOVA			non-parametric comparison				
			Chi-squared	df	p-value	estimator	lower	upper	statistic	p-value
		mono, multi	77.606	2	< 0.001	0.668	0.535	0.779	2.961	0.009
		mono, o60				0.125	0.07	0.214	-7.136	< 0.001
		multi, o60				0.075	0.034	0.157	-7.096	< 0.001

Table 4.7: ANOVA and post hoc test of absolute (VOT) and normalized (pVOT) voice onset time

As listed in Table 4.6, the plosive count for monosyllabic target words with a word-initial unaspirated alveolar stop [t] is considerably higher with 68 in the *mono* group, 48 stops in the *multi* group, and lastly 86 alveolar stops in the *o60* group.

The older Zurich German speakers have considerable lower mean VOT of 19 ms and median VOT of 17 ms, and also display the lowest variability with a standard deviation of 10 and median absolute deviation of nearly 6. The measured VOT in the multiethnolect group last around two and half times as long with a mean VOT of 49 ms and a median VOT of nearly 52 ms. The variability is also rather high with a SD of 20.5 and MAD of nearly 21. The *mono* group has a mean of 38 ms and median VOT of 35 ms with a variability of nearly 17 in standard deviation and 13 in MAD.

The *o60* group also displays by far the lowest pVOT measurements with a mean of 0.08 and median of 0.06 with variability measures of 0.05 SD and 0.04 MAD. By comparison, the multiethnolect speakers display a noticeable higher mean pVOT of 0.24 and median pVOT of 0.23 with identical variability measures of 0.12. The *mono* group is set in between with a mean pVOT of 0.19 and median pVOT of 0.17, and a standard deviation of 0.09 and MAD of 0.08.

The non-parametric ANOVA is significant for both pVOT and VOT for alveolar stops. The post hoc tests in Table 4.7 show the same significant results of VOT and pVOT in regard of generational comparisons with a p-value of < 0.001. However, there is a discrepancy in the results of the comparison between the younger groups where the p-value of VOT with 0.009 is clearly significant, but the p-value of 0.064 in the pVOT comparison is just above the significance level of 0.05.

5 Discussion

5.1 Comparison between groups and place of articulation

The distribution curves of VOT measurements of the general comparison of unaspirated stops (Figure 4.1a) as well as the reported statistical calculations in Section 4.1 meet our expectations. The distribution curve of the control group is clearly different from both adolescent groups. The speakers of the *o60* group show overall the shortest VOT durations with the lowest variability. The low variability indicates that the speakers produce VOT durations consistently short. As anticipated, the multiethnolect speakers show the longest VOT durations and highest variability. Their measurements are the most different from the control group. The *mono* speakers are positioned between the other two groups but seem to be more similar to their peers of the *multi* group since the VOT durations, medians, and the variability values of the two groups of younger speakers are rather close together, but distinct from the *o60* group. The statistical analysis additionally confirmed that all three groups produce distinct aspiration.

From these first findings, we can conclude that there is a clear generational difference in the use of aspiration by Zurich German speakers. Additionally, there is a difference in VOT production among the two adolescent groups. These assumptions are based on all unaspirated fortis plosives in the DiaPix corpus regardless of place of articulation and aspiration. The higher variability combined with the overall longer VOT measurements reinforce the observation of increase in aspiration in fortis plosives among younger speakers by different sources [Schifferle, 2010, 11; Ladd and Schmid, 2018, 232, 246; Leeman et al., 2020, 64].

When differentiating the stops in the corpus according to the place of articulation, the statistical analysis of the comparison per groups of the alveolar stops /t/ show the same results as the above discussed comparison of all plosives in the corpus in terms of VOT duration: All groups produce distinctly different aspiration patterns in alveolar stops. The *o60* group show the lowest VOT, the multiethnolect speakers show again the longest VOT measurements, and the measurements of the *mono* group is set between the two groups. These results echo the findings from the first paragraph since the complete corpus consists predominantly of unaspirated alveolar stops in each case.

The comparison regarding all bilabial stops /p/ in the corpus show the same generational differences in VOT production. However, the post hoc test indicates that the younger speakers in both groups produce similar voice onset time in bilabial plosives. Possible explanations for this will be discussed in the following Section 5.2.

In summary, this first general comparison and statistical analysis of all plosives clearly establish a generational difference in the production in aspiration in both places of articulation. These findings therefore point to a sound change.

5.2 Differences in [±Aspiration]

This section splits the analysis further into aspirated and unaspirated stops and discusses the results of each group one after the other starting with the *o60* group before discussing general observations about aspiration and sound change.

The median VOT values for both aspirated bilabial and alveolar stops in the *o60* group are significantly higher compared to the medians of stops in the same respective place of articulation that have been classified as unaspirated. This demonstrates that older Zurich German speakers clearly produce distinct unaspirated and aspirated versions of both the bilabial and alveolar stops which was further statistically confirmed by the pairwise Wilcoxon comparison.

The distinction in aspiration is not as obvious for both adolescent groups. The *multi* group shows a higher median and variability in unaspirated bilabial stops, as shown in Figures 4.5a and 4.5b. Mean VOT measurements in the same dataset, however, show the inverse as the mean VOT is smaller in unaspirated stops [p] compared to the mean measurements for aspirated bilabial stops [p^h]. The plosive count in this comparison is rather low with merely 13 unaspirate bilabial stops. An inspection of the individual VOT measurements of these 13 stops reveal a lot of variability. Speaker *Le35* produced *pèèrli* 'couple' with a voice onset time of only 16 ms. On the other hand, the VOT of *Bu22* in *pèèrli* or *pèèrlis* ranges from 31 to 54 ms in 5 utterances. Similar variability can be found in the stops that are categorized as [+Aspiration] where the VOT by speaker *Bu22* for *pinke* and *pinki* ranges from 24 to 51 ms, or a more extreme example is participant *Bu17* whose VOT in the stops *pink* and *pinke* last from 52 to 83 ms. The small token count combined with the high variability makes it difficult to formulate a general statement about the aspiration pattern of multiethnolect speakers. The additional Wilcoxon comparison within the *multi* group showed that there is not enough evidence in the corpus to substantiate that multiethnolects produce unaspirated and aspirated bilabial stops in a significantly different manner.

In consequence, the multiethnolect speakers do not seem to produce distinct VOT in bilabial stops, also considering the facts that the median VOT for both unaspirated (42 ms) and aspirated bilabial stops (37 ms) are not very far apart, and they show the identical variability (14.83).

The interpretation of measurements in aspirated alveolar stops [t^h] is based on a small token count of 31 stops in all three groups combined. The comparison is more or less dominated by one word, namely *tiischört* 't-shirt', or its shorter version *tiischi* (as well as written variations *tischört* and *tischi*) which make up for 22 of the 31 aspirated stops in the comparison.

This also applies to aspirated alveolar stops produced by the *multi* group. 6 of the 9 aspirated alveolar stops are variations of *tiischört*. The remaining three tokens are the target word *täil* 'object'[1]. There are two exceptionally long alveolar stops by multiethnolectal speakers, the first lasts 66 ms (*tiischört* by speaker *Le29*) and the other 75 ms (*täil*) by *Le22*). Variability can be found in only one speaker *Le31* that produces *täil* once with a VOT of 15 ms and the second time with 36 ms. However, these observations are based on a handful of alveolar stops which are probably too few examples to draw conclusions about the whole multiethnolect group.

In consequence, the observation that multiethnolect speakers generally produce shorter VOT in alveolar stops that are expected to be aspirated compared to alveolar stops that not expected be aspirated should be taken with caution.

Nevertheless, the Wilcoxon comparison in Table 4.5 suggests that multiethnolect speakers do not produce significantly different unaspirated and aspirated alveolar fortis plosives. As a result, it can be concluded that, in contrast to the *o60* group who produces different VOT patterns depending on the categorization in terms of aspiration, the speakers of the *multi* group produce similar aspiration patterns in stops that are categorized as unaspirated as well as aspirated regardless of place of articulation.

Against expectations, the speakers of the *mono* group show the longest VOT measurements in bilabial fortis plosives that are categorized as unaspirated [p] as shown in Figure 4.5b. The median VOT of 51 ms is significantly higher compared to a median of 42 ms of aspirated bilabial stops. This surprising result however is based on merely 7 unaspirated plosive. This is the smallest dataset in the complete analysis and the question arises how representative these are. A closer inspection reveals that the target word types are *puppe, pudel, passt, pèèrli*, and *putzmittel* 'doll, poodle, [it] fits, couple, cleaning products' (*pudel* appears three times). These 7 bilabial stops are uttered by only 4 of the 10 speakers of the *mono* group (*Le07, Le08, Le10,* and *Le33*). Interference from Standard German could play a significant role in these specific target words as most of them are clearly aspirated in Standard German [Morand et al., 2021]. With all these factors in mind, the question arises how representative these measurements are. It is certainly premature to deduce that *mono* speakers produce longer aspiration in stops that are classified as unaspirated. It is much more likely that the compiled data is not representative of the production of unaspirated bilabial stops by *mono* speakers.

The Wilcoxon test further showed that the *mono* group does not produce significantly different VOT in aspirated and unaspirated bilabial stops as the test showed a p-value of 0.32 for the data sets of [p] compared to the data set of [p^h]. These results might be interpreted as a first hint that *mono* speakers may produce unaspirated and aspirated bilabial stops in the same manner. Considering the above discussion on the data set, it is difficult to assert if *mono* speakers differentiate between aspirated and unaspirated

[1] *täil* is the only minimal pair in the corpus and would therefore be an optimal candidate for a comparison but unfortunately, there is no speaker in the corpus that produces both [tæi̯l] 'component' and [t^{h}æi̯l] 'object'.

bilabial stops based on this data alone, and further investigations is necessary to provide more solid evidence.

In terms of alveolar stops that are expected to be aspirated [t^h], the 6 *mono* speakers produce similar results as the *o60* group by producing distinct longer VOT. Though there are only 12 stops to support this claim which are based on solely two target word types: *täil* and *tischört* (and variations of it). This indicates that younger more traditional Zurich German speakers of the *mono* group have the same aspiration pattern as the speakers of the older generation in terms of aspirated alveolar stops.

In contrast to bilabial stops, the Wilcoxon comparison of alveolar stops shows that the speakers of the *mono* group produce significantly different aspiration patterns in alveolar stops, i.e. [t] and [t^h] compared to [p^h]. This contrariness confirms that more investigation is needed to establish if there is a significant difference in aspirated and unaspirated fortis plosives in the *mono* group.

The results show that the *o60* group clearly displays a three-way contrast in stop consonants in both places of articulation ([b̥ p p^h]). Based on the statistical calculations in Table 4.5, multiethnolect speakers do not differentiate between aspirated and unaspirated stops regardless of place of articulation and therefore seem to have only a two-way stop contrast ([b̥ p^h]).

For the adolescents who speak a more traditionally perceived Zurich German the calculations show different results based on the place of articulation of the stop consonant. Whereas they seem to display a three-way contrast for alveolar plosives ([d̥ t t^h]) and show, as expected, longer VOT measurements in aspirated alveolar stops, they lean towards a two-way contrast for bilabial stops ([b̥ p^h]). However as previously discussed, the data for unaspirated bilabial stops by the *mono* group might not be representative to draw conclusions on aspiration patterns. It is probably more plausible that the *mono* group displays the same three-way contrast in both places of articulation.

As discussed above, there are considerable indications for a generational difference in the production of VOT in fortis plosives. The question is, however, if the sound change is categorically based on increasing aspiration of lexical items [Ladd and Schmid, 2018, 247] or gradual as Morand et al. [2021, 15-16] observe an increasing aspiration of fortis plosives in multiethnolect speakers, as well as in speakers that are perceived to speak a more traditional Zurich German in a study on phonetic interference in Zurich German. As Leeman et al. [2020, 62] and Morand et al. [2021, 15-16] state, this may be explained by an adaption of Standard German where fortis plosives are indeed aspirated unlike the typically unaspirated fortis plosives in Zurich German.

The question of categorical or gradual sound change is challenging to answer based on the comparisons made in this study alone. The findings of this study seem to point to a gradual increase in aspiration among younger generations because our findings showed overall longer VOT measurements in younger speakers compared to older Zurich German speaker in unaspirated alveolar stops. However, this study did no comparison on specific lexical items to study a categorical sound change. Ladd and Schmid [2018, 247] observe that an increase of aspiration is primarily affecting low-frequency and obvious loan words. As reported in Table 7.1 in the appendix, most target words with word-initial fortis plosives are infrequent in the corpus which means they are hardly candidates in the course of this study since there would be too few, if any for that matter, tokens in individual groups. The five most frequent target word types with a word-initial unaspirated stop in the corpus are all alveolar stops: *tisch* 'table' (89 occurrences), *tüüre* 'door' (55 occurrences), *tüür* 'door' (47 occurrences), *tiger* 'tiger' (20 occurrences), and *tuech* 'cloth' (19 occurrences). All remaining target word types with a word-initial unaspirated plosive occur 15 times or rarer. The possibility to compare target words with a word-initial unaspirated bilabial stop is limited in this corpus, e.g. the most frequent word is *pudel* 'poodle' with 7 occurrences by the speakers of the *mono* and *o60* groups, but no occurrences in the *multi* group. The discussion is therefore limited to alveolar stops. The five most frequent target word types are not obvious loan words and therefore disagree with the reasoning by Ladd and Schmid [2018]. Additionally, all of them are unaspirated alveolar stops. As shown before, our findings suggest that all three groups produce distinctly different VOT durations in unaspirated alveolar stops, and therefore point to a gradual sound change.

Ladd and Schmid [2018, 247] observe that word-initial bilabial stops are produced more variable in terms of aspiration compared to word-initial alveolar stops. They also state that these differences can probably be traced back to the fact that most analyzed words with a word-initial bilabial stop are rare or loan words. When looking at unaspirated bilabial stops [p] and alveolar stops [t], the *mono* and *o60* groups both show a higher variability in word-initial bilabial stops as well, though the difference between median and MAD is more obvious in the *mono* group. The MAD of the multiethnolect speakers, however, show the opposite and are higher for alveolar stops. Of the 106 target word types in Table 7.1 around a third (39) have a word-initial bilabial stop. 36 types occur 7 times or less in the corpus. The three most frequent types are variations of the word 'pink' (*pink, pinke, pinki*) which is foreign and marked as aspirated. With a total token count of 50, the word 'pink' and its variations make up nearly half of all word-initial bilabial stops. About half (20) of the 39 bilabial types are loanwords and the total token count is 73. In summary, the data seems to reflect the observations made by Ladd and Schmid [2018, 247], as more than half of the bilabial stop are loanwords and most of them have a low occurrence count, except for the three variations of 'pink'. However, the assumptions are based on a very small token count in this corpus and should therefore be taken with caution.

5.3 Gender

There is a slight bias towards female speakers, since 18 of the 30 recorded speakers are female. Whereas the distribution of gender is relatively equal in multiethnolect and older speakers, there are only 2 male but 8 female speakers in the *mono* group. This circumstance should be considered in every comparison that includes male speakers of the *mono* group.

The gender analysis regarding alveolar stops does not differentiate plosives in terms of aspiration because the data set for aspirated alveolar stops would be extremely small since approximately 92% of all 407 alveolar stops in the corpus are unaspirated. The results for gender comparison therefore mostly reflect the result of unaspirated alveolar fortis plosives. The results of alveolar stops per group are discussed first.

The comparison of alveolar plosives uttered by female and male speakers show a clear generational difference in VOT production. These findings display the same outcome as the comparison of unaspirated alveolar stops [t], as well as the comparison between all alveolar stops /t/. The pairwise comparison within each group, as shown in Table 4.5, shows no differences in VOT duration between female and male speakers in any group. In conclusion, female and male speakers of all analyzed groups do not seem to produce VOT differently in alveolar stops.

The outcome of the bilabial stop comparisons are not as straightforward. On the one hand, there seems to be a clear generational difference in the comparisons of male speakers. The female speakers however seem to produce the same VOT duration between all groups since there is no significant result between groups in any comparison for bilabial stops. But every plosive count per group is rather small.

The pairwise comparison between female and male speakers for bilabial stops in the *mono* group is significant (p-value = 0.044), but very close to the significance level of 0.05. Statistically speaking the result is clear. But considering the small set of compared data, the question arises how representative this statistical result is. As seen in Figure 8.18 in the appendix, their distributions are quite distinct from each other. However, the comparison is dominated by aspirated stops that make up 85% of all 48 bilabial stops in the *mono* group. Therefore, there is a bias towards aspirated stops and female speakers in this particular comparison. In consequence, further testing is necessary to assert if there is a significant difference between female and male speakers in the *mono* group.

The *multi* group does not show a significant difference between gender as shown in Table 4.5, as well as in Figure 8.19 in the appendix. Additionally, both female and male speakers have the same plosive count. On this basis, the multiethnolect speakers seem to produce similar VOT durations regardless of gender.

The pairwise Wilcoxon comparison of female and male speakers of the *o60* group is clearly significant. As the *mono* group above, the inspection of the plosives per group showed that all 9 bilabial stops uttered by male speakers are unaspirated, which also explains the low median of 13 ms. The 24 plosives spoken by female speakers contain more than half of aspirated bilabial stops which is reflected in the median of 34 ms. In conclusion, these findings are probably not enough evidence to assert if there is an actual difference between gender within the *o60* group and more tests are necessary that compare aspirated and unaspirated bilabial stops separately.

In summary, the data does not provide substantial evidence to confidently assert gender differences in VOT production in the analyzed groups.

5.4 Speech rate normalization

The reported VOT measurements in Section 4.6 are all in absolute numbers which does not take individual speech rate into account. Morand et al. [2020a, 566] observe that adolescents, who are perceived as more ethnolect, speak slower compared to adolescents of the *mono* group. In consequence, speech rate may play a significant role in the comparison of VOT. As described in Section 3.6, pVOT is the ratio of the voice onset time compared to the duration of the remaining word as applied by Kleber [2018, 472] who compared minimal pairs by younger and older speakers of the German Bavarian and Saxon dialects. The pVOT ratio is only applicable for words of same length, or similar number of segments. The pVOT of a longer word such as *teddibäärlade* 'teddybear store' is always distinctly smaller compared to the pVOT of a shorter word as in *topf* 'pot' regardless of speech rate. The comparison is meaningless in this example since we cannot draw an informative conclusion if the speech rate has an influence on VOT. The analysis for speech rate normalization is therefore only applied to words with equal syllable length.

The reported results in Section 4.6 show the same outcome for absolute (VOT) and normalized (pVOT) measurements in monosyllabic words with a word-initial bilabial stop. Even though the compared data sets are rather small, these findings confirm the same observations made in Section 5.2 about aspirated bilabial stops. Hence, regardless of speech rate, all compared groups seem to produce VOT in aspirated bilabial stops significantly different from one another.

The calculations for absolute (VOT) and normalized (pVOT) measurements for monosyllabic words with a word-initial alveolar stop confirm a clear generational difference in VOT duration[2]. However, the pVOT comparison between the *mono* and *multi* groups is not significant and contradicts the observations made in Section 5.2. This suggests that, considering speech rate, the younger speakers in this study may produce same VOT durations after all. However, the p-value that this conclusion is based on is just barely above the significance level. Mathematically speaking, there is no difference between the compared groups for unaspirated alveolar stop. Considering the data and observations made above, this assumption must be investigated more thoroughly in future work by comparing the same target words and not just words with similar segment counts to confidently assert if there is a significant difference between the two groups of younger speakers.

5.5 Speaker variability

Schifferle [2010, 43] reports variability within individual speakers who pronounce the same word sometimes with and sometimes without aspiration, as in *Tämperatuure* vs. *Thämper-*

[2] An additional comparison for two-syllable words with a word-initial unaspirated alveolar was made, but the results (as listed in Tables 7.2 and 7.3 in the appendix) are not reported in detail because they show the same outcome as unaspirated alveolar in monosyllabic words. They reinforce however the outcome of the monosyllabic comparison.

atuure 'temperature'. The same variability most probably occurs frequently in this corpus as selectively reported in Section 5.2. The following paragraphs give an insight in the variability of VOT durations of the individual speaker of each group. Even though no statistical analysis was made on this subject, some differences between the groups are discussed on the basis of Figure 5.1 that displays individual speaker variability in each group. The corresponding mean, median, standard deviation, and median absolute deviation of VOT measurements for each of the 30 participants are listed in Table 7.4 in the appendix.

As shown in Figure 5.1, the VOT distribution curves of the older Zurich German speakers are rather similar which reflects the low combined variability MAD score of 7.41 in the *o60* group (listed in Table 4.2). Speakers *Ze06*, *Ze07*, and *Ze10* stand out for having similar and more shallow distribution curves. However, the similarity in shapes and the reported MAD values suggest that all speakers of the *o60* group produce stops with more consistent VOT durations.

The distribution curve of *Le05* stands out among the other speakers in the *mono* group. The sharp shape is very reminiscent to the curve of most speakers in the *o60* group but shifted to the right with a higher median VOT compared to the *o60* group. The second peak on the right side in the curve of *Le05* is most likely an outlier. The remaining distribution curves can be split roughly into two groups. The first one has higher peaks that suggest lower variability with shorter VOT durations. The second with more shallow shapes indicate more variability with overall longer voice onset time.

The *multi* group shows similar distribution curves to the *mono* group in terms of height which indicates that all younger speakers in both groups display similar variability. The peaks of the curves are however positioned in higher VOT durations which shows that multiethnolect speakers generally produce longer voice onset time than the *mono* group. The rightmost distribution curve belongs to speaker *Bu18* who displays the highest screening score and therefore was perceived as the most multiethnolect of all adolescents. This may suggest that a high screening score is connected to a long VOT production.

In summary, the density graphs per group visualize clearly the generational difference in terms of VOT duration where older Zurich German speakers show shorter VOT, i.e. overall less aspiration in fortis plosives, and the younger speakers of both younger groups tend to produce longer VOT, i.e. overall more aspiration in fortis plosives. Additionally, the distinct consistency in VOT production by the *o60* group contrasts the variability found in both younger groups.

5.6 Limitations

The categorization of stops consonants in the corpus in terms of [±Aspiration], as displayed in Table 7.1 was done by the author of this study – a young native Zurich German speaker. The classification in terms of aspiration is the basis of this analysis and might therefore be biased to one perspective. Whereas the classification in terms of [±Aspiration] might be clear in some cases, other words with word-initial stops leave more room for debate as

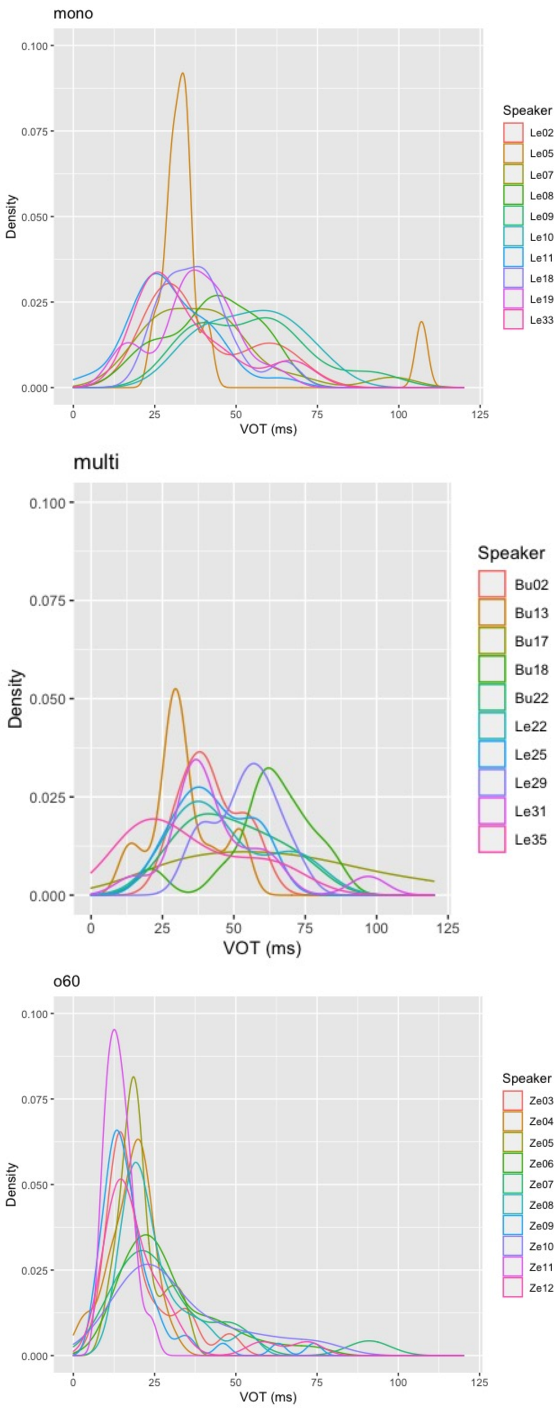


Figure 5.1: Speaker variability in VOT per group

they heavily vary among questioned speakers in the peergroup of the author, for example *türkis* 'turquoise', or the verb *picked* '[they] pecked' which is marked as [+Aspiration]. Additionally, there is no inter-annotator agreement since the annotation was performed by only by the author as well. In consequence, all measurements, categorization, as well as results and findings are based on the assessment of a single perspective.

The fortis plosive corpus of this study displays some limitations. On the one hand, the spontaneous speech of the recordings might reflect the use of fortis plosives in everyday speech of all speakers more accurately. On the other hand, the greedy approach in plosive selection to gather as much data as possible resulted in a highly variable corpus which shows limitations. Especially in the sense that there is a highly varying number of tokens per speaker and not every speaker produces the same target words which are uttered in different phonetic contexts. This also complicates the comparison of speech rate.

Further, the number of tokens per group comparison also heavily differs which leads to some rather small amount of stops that are compared, particularly for the group comparison of unaspirated bilabial stops [p] and aspirated alveolar plosives [t^h], as previously discussed. The analysis of aspirated plosives [p^h] and [t^h] are also heavily dominated by basically one word each: *pink* and *tiischört* (including their written variations). The corpus displays an additional imbalance in token size per group. Almost half the tokens were produced by older Zurich German speakers, and only about a fifth by the *multi* group. 119 of all analyzed stops are bilabial compared to 407 alveolar plosives. Therefore, further studies with a planned selection of fortis plosives are needed to bring evidence to the observations made in this study.

24 of all 526 analyzed plosives do not place the stress on the first syllable and could be excluded due to the small amount. They were however included, because most of them are bilabial stops (16 of 24 tokens) and excluding them would mean that some speakers would have fewer or no bilabial stops in their data sets. 8 of them are variations of *türkis* 'turquoise', which in standard German is clearly stressed on the second syllable, in Zurich German however highly variable depending on speaker. In further investigations, the selection of plosives should be more uniform.

This analysis was conducted with absolute numbers and worked closely with the 'raw' data that does not take speech rate into account. Conclusions rely on the calculations of the statistical analysis that were chosen on the basis of the textbook on how to do linguistic analysis in R [Levshina, 2015]. The chosen statistical tests show two inconsistencies with the results of the post hoc test. These discrepancies regard the significant results between groups comparisons for unaspirated alveolar stops [t^h] and for bilabial stops by female speakers, which means there should be a difference between the three groups. The post hoc tests for both comparisons however show no difference in VOT production. It is difficult to identify the reason for these inconsistencies. The small plosive count per group or the methods of the statistical tests could be a possible explanation for this outcome.

Either way, the findings of this study should be taken with caution especially due to the small number of samples in some comparisons. Nevertheless, the findings of this empirical study clearly confirm an increase in aspiration in the younger generation of Zurich German speakers. Furthermore, they demonstrate that further investigations in the terms of a sound change is reasonable.

6 Conclusion

An increase in aspiration in fortis plosives among younger Swiss German speakers has been observed, among others, by Schifferle [2010, 11]. The generational difference in the use of aspiration in stop consonants may indicate a sound change. Yet, no empirical study has been conducted on the aspiration in fortis plosives in Zurich German, but there has been observations on the topic of sound change in Zurich German. Ladd and Schmid [2018, 246-247] discussed a possible categorical sound change in the aspiration patterns of Zurich German stop consonants through lexical diffusion. Morand et al. [2021, 14-15] however observe a general increase in aspiration in fortis plosives by adolescent multiethnolect Zurich German speakers, as well as adolescents that are perceived to speak a more traditional Zurich German. These findings suggest a more gradual sound change.

This study aims to analyze generational differences by applying the *apparent time paradigm*. The method assumes that the way of speaking of older generations does not change since their adolescence. Therefore, older speakers display the language before a suspected sound change. Thus, data of older and younger speakers of Zurich German are compared. The investigation additionally compares multiethnolect Zurich German speakers. Multiethnolect is a relatively new way of speaking that has emerged from multiethnic neighborhoods in multiple European cities, including Zurich city.

In consequence, VOT measurements of three groups, each containing 10 Zurich German speakers, are compared. The participants of the *o60* group are over 60 years old and speak traditional Zurich German. The remaining adolescent speakers are categorized into the groups *mono* if they are perceived as speaking a more traditional Zurich German by peers, or *multi* if their speech is perceived as more multiethnolect.

We expected that older Zurich German speakers show the shortest VOT measurements, and in contrast that the multiethnolect speakers show the longest VOT measurements. The VOT measurements of speakers of the *mono* group should be positioned between the other two groups. We additionally investigated if there is a difference in VOT durations between female and male speakers within each group.

The VOT measurements of Zurich German bilabial and alveolar fortis plosives were extracted from the DiaPix corpus of the *Phonetic features of (multi-)ethnic urban vernaculars in German-speaking Switzerland* project. 526 stop consonants were found in the spontaneous speech recordings using Praat scripts and regular expressions. The web application *Webmaus* was used for the automatic phonetic annotation in TextGrid files. Additional Praat scripts were used to alter the TextGrid files according to the needs of the study. All

VOT measurements were manually segmented. The statistical analysis was conducted by applying variations of non-parametric ANOVA [Levshina, 2015].

The analysis is split in several smaller comparisons. The main conclusion of the VOT analysis is a clear increase in aspiration by younger Zurich German speakers compared to older Zurich German speakers. This is substantiated in several comparisons but made especially abundantly clear in the comparison of the 376 unaspirated alveolar stops which is based on a relatively large data set. The differences between the two groups of younger speakers is not as obvious in all comparisons.

The comparison of all alveolar stops clearly show differences in VOT durational between all groups. Regarding all bilabial stops, we could only determine a distinct difference in VOT length between generations.

Additionally, separate analyses of fortis plosives in terms of place of articulation and aspiration was conducted. The results suggest that all Zurich German speakers produce similar VOT durations in fortis plosives that are expected to be aspirated. However, the analysis of traditionally unaspirated plosives is more insightful to determine sound change differences. The statistical analysis demonstrates that there is a significant difference in VOT measurements for unaspirated alveolar plosives between all groups. These differences are visualized in Figure 4.5d and reflect the expected outcome in terms of VOT differences. The *o60* group shows the shortest VOT, the *multi* group the longest, and the VOT measurements of the *mono* group is situated between them. The analysis of unaspirated bilabial plosives further confirms the generational difference in aspiration duration as shown in Figure 4.5b. The unexpected overly long VOT measurements of the *mono* group in unaspirated bilabial stops are most likely not representative due to a very small dataset. Additionally, the target words in the data set do not seem to be optimal candidates for the comparison of unaspirated bilabial stops.

Further, the fortis plosive corpus did not provide substantial evidence to show aspiration differences between female and male speakers in any group.

Within group comparisons revealed that older Zurich German speakers clearly differentiate between aspirated and unaspirated stops and therefore show a three-way contrast in stop consonants regardless of place of articulation ([b̥ p p^{h}]). They even produce considerably longer VOT durations in aspirated stops compared to both groups of younger speakers. The VOT measurements of multiethnolect speakers showed no significant difference between aspirated and unaspirated stops in both places of articulation and in consequence show a two-way stop contrast ([b̥ p^{h}]). The *mono* group showed a clear three-way contrast in the analysis of alveolar stops ([d̥ t t^{h}]), but a two-way contrast in bilabial stops ([b̥ p^{h}]). The small data set of bilabial stops however might no be representative and it is plausible to assume that the speakers of the *mono* group produce a three-way contrast in both places of articulation. Still, further analysis is necessary to determine the types of plosives that are produced by the *mono* group.

The above discussed findings are based on absolute numbers taken from spontaneous speech recordings. The comparisons with the normalized voice onset time ratio pVOT reinforce the clear generational difference in aspiration in monosyllabic and two-syllable words. However, the normalized pVOT showed no significant difference between both groups of younger Zurich German speakers. Further investigation is therefore necessary to establish if there is a significant difference in aspiration between the *mono* and *multi* groups.

A visual inspection of individual speakers per group revealed that older Zurich German speakers show the shortest VOT measurements and less variability compared to the multiethnolect speakers who demonstrate the longest VOT durations and the most variability within the group. The speakers of the *mono* group is again positioned between the other groups in terms of variability and VOT duration.

The corpus and the conducted analysis show some limitations. Many parameters of the study on which the findings are based depend on one single perspective in the course of the master thesis, inherently the classification if a fortis plosive is traditionally aspirated or unaspirated, or the annotation itself since there was no inter-annotator agreement. Additionally, the fortis plosives were extracted from recordings that were not intended to elicit Zurich German fortis plosives. As a result, there is a varying number of fortis plosives and target word types per speakers. Furthermore, the participants uttered different target words in highly different contexts. Also, the token count per compared group varies greatly.

Having said that, the findings of this study should be taken with caution and not interpreted as evidence but as indications. Nevertheless, the findings of this empirical study on bilabial and alveolar fortis plosives clearly show an increase in aspiration among the younger generation of Zurich German speakers. Furthermore, the findings suggest that a gradual sound change is in progress that is more prominently apparent in the multiethnolect speakers. The reason for the sound change will not be discussed in this study. It is unclear at this stage if the sound change emerged because of multiethnolect speakers. However, their aspiration pattern might fuel it. The conclusions illustrate that further investigation in aspiration differences among younger Zurich German speakers is necessary and justified.

References

Aho, A. V. (1991). *Algorithms for finding patterns in strings, Handbook of theoretical computer science (vol. A): Algorithms and complexity.* The MIT Press, Cambridge, Massachusetts.

Auer, P. (2003). 'Türkenslang': ein jugendsprachlicher Ethnolekt des Deutschen und seine Transformationen. In Häcki-Buhofer, A., editor, *Spracherwerb und Lebensalter*, pages 255–264. Francke Verlag, Tübingen & Basel.

Baker, R. and Hazan, V. (2011). DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior Research Methods*, 43(3):761–770. Springer.

Beckman, J., Jessen, M., and Ringen, C. (2013). Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics*, 49(2):259–284. Cambridge University Press.

Boersma, P. and Van Heuven, V. (2001). Speak and unSpeak with PRAAT. *Glot International*, 5(9/10):341–347.

Braun, A. (1988). *Zum Merkmal 'Fortis/Lenis': phonologische Betrachtungen und instrumentalphonetische Untersuchungen an einem mittelhessischen Dialekt.* Number 55. Steiner, Stuttgart.

Bruno, M. (2019). Morphosyntaktische Merkmale von multiethnolektalem Zürichdeutsch. Master Thesis, University of Zurich.

Cheshire, J., Fox, S., Kerswill, P., and Torgersen, E. (2008). Ethnicity, friendship network and social practices as the motor of dialect change: Linguistic innovation in London. *Sociolinguistica*, 22(1):1–23. De Gruyter.

Deterding, D. and Nolan, F. (2007). Aspiration and voicing of Chinese and English plosives. In *Proceedings of the 16th International Congress of Phonetic Sciences*, pages 385–388. Universität des Saarlandes Saarbrücken Germany.

Enstrom, D. H. and Spörri-Bütler, S. (1981). A voice onset time analysis of initial Swiss-German stops. *Folia Phoniatrica et Logopeaedia*, 33(6):137–150.

Fagyal, Z. and Torgersen, E. (2018). Prosodic rhythm, cultural background, and interaction in adolescent urban vernaculars in Paris: case studies and comparisons. *Journal of French Language Studies*, 28(2):165–179. Cambridge University Press.

Fleischer, J. and Schmid, S. (2006). Zurich German. *Journal of the International Phonetic Association*, 36(2):243–253. Cambridge University Press.

Fulop, S. A. (1994). Acoustic correlates of the fortis/lenis contrast in Swiss German plosives. *Calgary Working Papers in Linguistics (CWPL)*, 16:55–64. University of Calgary.

Hothorn, T., Hornik, K., van de Wiel, M. A., Zeileis, A., and Hothorn, M. T. (2019). Package 'coin'. *Conditional Inference Procedures in a Permutation Test Framework*.

Jakobson, R., Fant, C. G. M., and Halle, M. (1972). *Preliminaries to speech analysis: The distinctive features and their correlates.* The MIT Press, Cambridge, Massachusetts.

Kisler, T., Schiel, F., and Sloetjes, H. (2012). Signal processing via web services: the use case WebMAUS. In *Digital Humanities Conference 2012*, pages 30–34, Hamburg, Germany.

Kleber, F. (2016). 7. Lautwandel. In Domahs, U. and Primus, B., editors, *Handbuch Laut, Gebärde, Buchstabe*, pages 125–143. De Gruyter, Berlin.

Kleber, F. (2018). VOT or quantity: what matters more for the voicing contrast in German regional varieties? Results from apparent-time analyses. *Journal of Phonetics*, 71:468–486. Elsevier.

Konietschke, F., Placzek, M., Schaarschmidt, F., and Hothorn, L. A. (2015). nparcomp: an R software package for nonparametric multiple comparisons and simultaneous confidence intervals. *Journal of Statistical Software*, 64(1):1–17.

Kraehenmann, A. (2001). Swiss German stops: geminates all over the word. *Phonology*, 18:109–145. Cambridge University Press.

Labov, W. (1994). *Principles of linguistic change. Vol. 1: Internal factors.* Wiley-Blackwell, Oxford.

Ladd, D. R. and Schmid, S. (2018). Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops. *Journal of Phonetics*, 71:229–248. Elsevier.

Leeman, A., Jeszenszky, P., Steiner, C., Messerli, J., and Studerus, M. (2020). Behauchung von Plosiven im Schweizerdeutschen: Ein Merkmal junger DialektsprecherInnen? *Abstracts. 16. Tagung zur Phonetik & Phonologie im deutschsprachigen Raum*, pages 64–65.

Levshina, N. (2015). *How to do linguistics with R.* John Benjamins Publishing Company, Amsterdam-Philadelphia.

Lisker, L. and Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3):384–422. Taylor & Francis.

Machač, P. and Skarnitzl, R. (2009). *Principles of phonetic segmentation.* Epocha, Praha.

Morand, M.-A., Bruno, M., Julmi, N., Schwab, S., and Schmid, S. (2019). The voicing of lenis plosives in Zurich German: a sociophonetic marker of (multi-) ethnolectal speech. In Calhoun, S., Escudero, P., Tabain, M., and Warren, P., editors, *19th International Congress of Phonetic Sciences, Melbourne, Australia, 5 August 2019 - 9 August 2019*, pages 1600–1604.

Morand, M.-A., Bruno, M., Julmi, N., Schwab, S., and Schmid, S. (2020a). Speech rhythm in multiethnolectal Zurich German. In *10th International Conference on Speech Prosody 2020, Tokyo, Japan, 2020 - 2020*, pages 566–570. ISCA.

Morand, M.-A., Schwab, S., and Schmid, S. (2020b). The perception of multiethnolectal Zurich German: A continuum rather than clear-cut categories. *Loquens*, 7(2):1–12.

Morand, M.-A., Schwab, S., and Schmid, S. (2021). Standarddeutsche Interferenzen im Dialektwortschatz Schweizer Jugendlicher: Lexikalische und lautliche Entlehnungen. pages 1–18. to appear in Bulletin VALS/ASLA 114.

Quist, P. and Svendsen, B. A., editors (2010). *Multilingual urban Scandinavia: New linguistic practices.* Multilingual matters, Bristol, Buffalo, Toronto.

R Core Development Team (2016). R: A language and environment for statistical computing, version 4.0.0.

Rampton, M. B. H. (1987). *Uses of English in a multilingual British peergroup.* PhD thesis, Institute of Education, University of London.

Renwick, M. and Ladd, D. (2016). Phonetic distinctiveness vs. lexical contrastiveness in non-robust phonemic contrasts. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1):1–29. Ubiquity Press.

Schifferle, H.-P. (2010). Zunehmende Behauchung. Aspirierte Plosive im modernen Schweizerdeutsch. In *Alemannische Dialektologie: Wege in die Zukunft. Beiträge zur 16. Tagung für alemannische Dialektologie in Freiburg/Fribourg vom 07.–10.09.2008*, pages 43–55. Steiner, Stuttgart.

Schmid, S. (2010). Segmental features of Swiss German ethnolects. In *Sociophonetics at the Crossroads of Speech Variation, Processing and Communication Conference, Scuola Normale Superiore Pisa*, pages 69–72.

Schmid, S. (2019a). Phonology Lecture 3. Lecture at the University of Zurich.

Schmid, S. (2019b). Phonology Lecture 7. Lecture at the University of Zurich.

Schmid, S. (2020). The Sounds of the World's Languages Lecture 6. Lecture at the University of Zurich.

Sievers, E. (1876). *Grundzüge der Lautphysiologie zur Einlührung in das Studium der Lautlehre der indogermanischen Sprachen.* Breitkopf und Hartel, Leipzig.

Tissot, F., Schmid, S., and Galliker, E. (2011). Ethnolektales Schweizerdeutsch. In Glaser, E., Ericht Schmidt, J., and Frey, N., editors, *Dynamik des Dialekts: Wandel und Variation*, pages 319–344. Steiner, Stuttgart.

Van Rossum, G. and Drake Jr, F. L. (1995). *Python tutorial.* Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands. version 3.9.2.

Weber, A. (1948). *Zürichdeutsche Grammatik: ein Wegweiser zur guten Mundart*, volume 1. Schweizer Spiegel Verlag, Zurich.

Winteler, J. (1876). *Die Kerenzer Mundart des Kantons Glarus.* CF Wintr'sche Verlagshandlung, Leipzig.

Würth, K. (2002). Zur phonologischen Klassifikation der Fortis/Lenis-Opposition im Zürichdeutschen. Master Thesis, University of Zurich.

Würth, K. (2020). *Consonant Quantity and Positional Neutralisation–Heusler's Law and Winteler's Law in Zurich German.* PhD thesis, University of Zurich.

7 Tables

Target word	Aspiration	Stress on syllable	Loan word	Syllable count	Occurence in corpus
tisch	-	+	no	1	89
tüüre	-	+	no	2	55
tüür	-	+	no	1	47
pink	+	+	yes	1	22
tiger	-	+	no	2	20
tuech	-	+	no	1	19
pinke	+	+	yes	2	18
tafle	-	+	no	2	15
tach	-	+	no	1	14
tier	-	+	no	1	12
pinki	+	+	yes	2	10
tüüren	-	+	no	2	10
täller	-	+	no	2	9
pudel	-	+	no	2	7
tiischört	+	+	yes	2	7
tischli	-	+	no	2	7
persoon	+	-	no	2	6
tischört	+	+	yes	2	6
paar	-	+	no	1	5
täfeli	-	+	no	3	5
täil_0	-	+	no	1	5
teddibäär	-	+	yes	3	5
tiischi	+	+	yes	2	5
pèèrlis	-	+	no	2	4
persoone	+	-	no	3	4
poster	+	+	yes	2	4
tischi	+	+	yes	2	4
toor	-	+	no	1	4
türkis	-	-	no	2	4
tüürflügel	-	+	no	3	4
pèèrli	-	+	no	2	3
pudl	-	+	no	1	3
täil_1	+	+	no	1	3
tedibèèr	-	+	yes	3	3
topf	-	+	no	1	3
päärli	-	+	no	2	2
picknick	+	+	yes	2	2
pisse	+	+	yes	2	2
puli	-	+	yes	1	2
putzmittel	-	+	no	3	2
tächlichappe	-	+	no	4	2
taflen	-	+	no	2	2
täili_1	+	+	no	2	2
teke	+	+	yes	2	2
tube	-	+	no	2	2
tuet	-	+	no	1	2
türe	-	+	no	2	2
türkiis	-	+	no	2	2
tüürraame	-	+	no	3	2
pälter	+	+	yes	2	1
papier	-	-	no	2	1
papiir	-	-	no	2	1
parallel	-	+	yes	3	1
parat	-	-	yes	2	1
passt	-	+	no	1	1
pèmpers	+	+	yes	2	1
petrol	-	-	no	2	1
picked	+	+	no	2	1
pinkes	+	+	yes	2	1
pinkgääls	+	+	yes	2	1
pinks	+	+	yes	1	1
pix	+	+	yes	1	1
pouster	+	+	yes	2	1
publikum	-	+	yes	3	1
pudeli	-	+	no	3	1
pul	+	+	yes	1	1
pullover	-	-	yes	3	1
pulover	-	-	yes	3	1
punkt	+	+	no	1	1
pünktlis	+	+	no	2	1
puppe	-	+	no	2	1
putze	-	+	no	2	1
tachabschluss	-	+	no	3	1
tachrand	-	+	no	2	1
täfli	-	+	no	2	1
tätigkäite	-	+	no	4	1
tätschle	-	+	no	2	1
tätschlet	-	+	no	2	1
teddibäärlade	-	+	yes	5	1
teddihuus	-	+	yes	3	1
tedibèèrli	-	+	yes	4	1
teken	+	+	yes	2	1
tèller	-	+	no	2	1
tigerchopf	-	+	no	3	1
tigerkopf	-	+	no	3	1
tochter	-	+	no	2	1
tone	-	+	no	2	1
töörffemer	-	+	no	3	1
top	+	+	yes	1	1
torte	-	+	no	2	1
tub	-	+	no	1	1
tuben	-	+	no	2	1
tues	-	+	no	1	1
tuesch	-	+	no	1	1
tünd	-	+	no	1	1
tünn	-	+	no	1	1
türkisblaue	-	-	no	4	1
türkise	-	-	no	3	1
turteschtuck	-	+	no	3	1
türm	-	+	no	1	1
turnschue	-	+	no	2	1
tut	-	+	no	1	1
tuube	-	+	no	2	1
tüübli	-	+	no	2	1
tüürgriff	-	+	no	2	1
tüüri	-	+	no	2	1

Table 7.1: Overview of all 106 types of target words and occurrences in corpus

Factor		Group	Plosive count	VOT Mean	VOT SD	VOT Median	VOT MAD	p-value Shapiro	p-value Bartlett	p-value Fligner	Normal Distribution	Homogeneity of Variance
[t]	pVOT	mono	31	0.13	0.06	0.14	0.05	0.133				
		multi	16	0.2	0.14	0.17	0.07	< 0.001		< 0.001	no	no
		o60	95	0.07	0.04	0.06	0.03	< 0.001				
	VOT	mono		33.16	10.97	32	11.86	0.395				
		multi		52.5	15.17	53.5	18.53	0.598		< 0.001	no	no
		o60		20	7.89	18	7.41	< 0.001				

Table 7.2: pVOT and VOT comparison of two syllable words

Factor		Group	Test between groups			Post hoc test				
[t]	pVOT		not parametric ANOVA			not parametric comparison				
			Chi-squared	df	p-value	estimator	lower	upper	statistic	p-value
		p(mono , multi)	49.179	2	< 0.001	0.698	0.475	0.855	2.099	0.093
		p(mono , o60)				0.152	0.081	0.266	-5.72	< 0.001
		p(multi , o60)				0.071	0.025	0.184	-5.578	< 0.001
	VOT		not parametric ANOVA			not parametric comparison				
			Chi-squared	df	p-value	estimator	lower	upper	statistic	p-value
		p(mono , multi)	78.484	2	< 0.001	0.85	0.657	0.944	3.767	< 0.001
		p(mono , o60)				0.155	0.082	0.273	-5.563	< 0.001
		p(multi , o60)				0.02	0.005	0.08	-6.392	< 0.001

Table 7.3: Statistical test to identify groups that are significantly different for two syllable words

Class	Speaker	#	Plosive	Aspiration	Annotation	Closure	VOT	Context
mono	Le02f	64	pink	+	p^h	0.135	0.053	pink
		65	pinke	+	p^h	0.064	0.035	und häts bi dir bi dem pinke huus igendwelchi sache a de tüüre
		66	pinke	+	p^h	0.037	0.061	und was gseesch du i dem èm pinke huus ine drin
		70	tedibèèr	-	t^h	0.078	0.069	näbet de apotek häts det so tedibèèr im schaufänschter
		71	tiischi	+	t^h	0.235	0.055	aso und dem mäitli mit dèm èm mit dem tiischi da
		72	tisch	-	t	0.078	0.022	häts bi dir uf dèm tisch aso dè blaui da rächts häts det au so èm en äimer mit wasser und eme schwamm
	Le05f	79	pink	+	p^h	0.071	0.029	s isch pink
		80	pinke	+	p^h	0.084	0.108	und bi dem pinke huus
		81	pinke	+	p^h	0.052	0.021	und händs dini mülltone hinderem pinke huus
		82	tisch	-	t	0.081	0.035	und denäbe häts nomal en tisch mit so drüü stüel
		83	tisch	-	t	0.06	0.031	aso ufem tisch
		86	tüüre	-	t	0.093	0.03	häts e tüüre
	Le08m	110	pink	+	p^h	0.2	0.046	pink
		111	pinkgääls	+	p^h	0.18	0.046	èm si hätt es pinkgääls
		112	Pudel	-	p^h	0.106	0.034	bi mir hätt si kän Pudel
		113	Täil_+	+	t^h	0.062	0.041	so s Täil i de Mitti
		114	Täil_+	+	t^h	0.022	0.062	ja das Täil wo oben abe luegt isch das orangsch wiiss
		115	Täller	-	t	0.089	0.033	uf mene blaue Täller
multi	Bu02m	1	paar	-	p	0.203	0.036	paar Sache am Bode
		2	pink	+	p^h	0.13	0.027	pink
		3	pinke	+	p^h	0.095	0.056	hätts bi dèm pinke Reschtorant oder Huus
		4	Teddibäär	-	t^h	0.03	0.036	hätt din Teddibäär so Schläiffe
		5	Tiger	-	t^h	0.073	0.04	wele Tiger

		6	Tiger	-	t^h	0.051	0.034	git kän Tiger
	Bu13m	11	Papier	-	p^h	0.024	0.054	es groosses Blatt Papier
		12	parallel	-	p	0.096	0.015	die sind wie parallel
		13	pink	+	p^h	0.163	0.022	schwarz pink
		16	Tach	-	d	0.067	0.003	weli Farb hätt de Tach
		17	Teddibäär	-	t	0.085	0.031	da nääbet de Drogerii hätt s zwäi Teddibäär gseesch die au
		19	Tiger	-	t^h	0.055	0.042	gseesch de Tiger uf em Plakat
	Bu18m	39	paar	-	b	0.056	0.002	säg maal du stell du paar Fraage
		40	pink	+	p^h	0.104	0.087	èm pink schwarz gschträifft
		41	Pul	+	p^h	0.054	0.057	mit en Menücharte uf em Pul èm uf em Tisch
		42	Tach	-	d	0.023	0.016	hätt s bi dier en Tach
		43	Täller	-	t^h	0.074	0.058	Kaffee st hätt s au so en Kueche druff zäichnet uf em Täller
		44	Teddibäär	-	t^h	0.019	0.059	wie gseend bi dir d Teddibäär us gseend bäidi gliich
o60	Ze03f	284	pèmpers	+	p^h	0.092	0.047	und rächts chunnt es beibi mit ere pèmpers und emene
		285	persoon	+	p	0.105	0.029	und emene chlinere persoon mit dunk dunkelbru-une haar
		286	pudel	-	p	0.065	0.033	und en pudel
		289	tächlichappe	-	t	0.08	0.014	vor dèm huus häts en maa blaue puli tächlichappe und bruune hose
		290	täller	-	t	0.097	0.014	und en blaue täller
		291	täller	-	t	0.042	0.014	ebe häts en täller und es stuck chueche bi mir
	Ze04m	302	papiir	-	p	0.1	0.005	bi der äinte linggs obe häts igendes es papiir wo öpis gschribe
		303	pudel	-	p	0.099	0.003	näi de pudel fèèlt
		304	pulover	-	p	0.099	0.01	ja aber das mit da s gschpängschtli isch nu das mit de roote hose das mit de blaue hose isch käs gschpängschtli sondern hät en gääle pulover und
		305	tächlichappe	-	t	0.072	0.013	und die tächlichappe hät e sonen wiisse
		306	tafle	-	t	0.233	0.01	e blaui tafle es hät e so es isch a uufghänkt
		307	täller	-	t	0.09	0.02	blaue täller miteme chueche druff wo dampft
	Ze05m	325	Päärli	-	p	0.084	0.013	und dänn es Päärli deet i de
		326	Päärli	-	p	0.113	0.017	und dänn de zwäiti Tisch mit em jüngere Päärli
		327	putze	-	p	0.076	0.024	han i es Chübeli und en Schwamm mit èm zur zum de Tisch putze hè
		328	Tachrand	-	t	0.132	0.032	und de Tachrand isch dunkelgrüen und relativ bräit
		329	Tafle	-	t	0.102	0.018	au und e Chette wo d Tafle zäme hebed so
		330	Tafle	-	t	0.176	0.025	und obe dra e Tafle

Table 7.5: Test corpus

	Speaker	Gender	Plosive count	VOT mean	VOT sd	VOT median	VOT mad
mono	Le02	f	15	39.53	15.64	32	7.41
	Le05	f	11	39.36	22.8	34	4.45
	Le07	f	18	40.78	19.41	38	14.83
	Le08	m	16	42.5	13.32	42.5	17.79
	Le09	f	22	54.82	17.49	56	20.76
	Le10	f	10	54.4	14.13	54.5	18.53
	Le11	f	25	30.24	12.6	28	10.38
	Le18	f	11	38.55	11.66	38	7.41
	Le19	m	28	37.18	12.74	37.5	10.38
	Le33	f	7	35.43	16.17	26	4.45
	Total		163	40.85	16.98	38	14.83
multi	Bu02	m	10	41.9	10.06	40	11.12
	Bu13	m	19	32.21	11.57	30	5.93
	Bu17	f	9	63.44	30.79	64	37.07
	Bu18	m	11	61.91	17.17	62	13.34
	Bu22	f	14	49.07	16.27	47	18.53
	Le22	f	7	47	17.6	41	11.86
	Le25	m	18	42.89	12.48	41.5	13.34
	Le29	m	11	53.91	11.35	54	11.86
	Le31	f	13	45.15	19.99	36	4.45
	Le35	m	8	34	20.13	27.5	18.53
	Total		120	46.03	18.83	42	17.79
o60	Ze03	f	18	19.83	10.4	15	2.97
	Ze04	m	23	17.52	6.58	19	4.45
	Ze05	m	24	20.38	6.49	19.5	5.19
	Ze06	m	31	29.45	15.02	26	10.38
	Ze07	f	14	30.71	21.09	25	11.12
	Ze08	f	15	24.93	10.74	21	5.93
	Ze09	f	48	18.75	13.04	15	5.93
	Ze10	f	28	32.07	18.43	26	12.6
	Ze11	m	21	13.95	3.89	14	4.45
	Ze12	f	21	22.1	15.59	16	4.45
	Total		243	22.72	14.14	18	7.41

Table 7.4: Speaker variability per group

8 Figures

8.1 Quantile-quantile and density plots per comparison

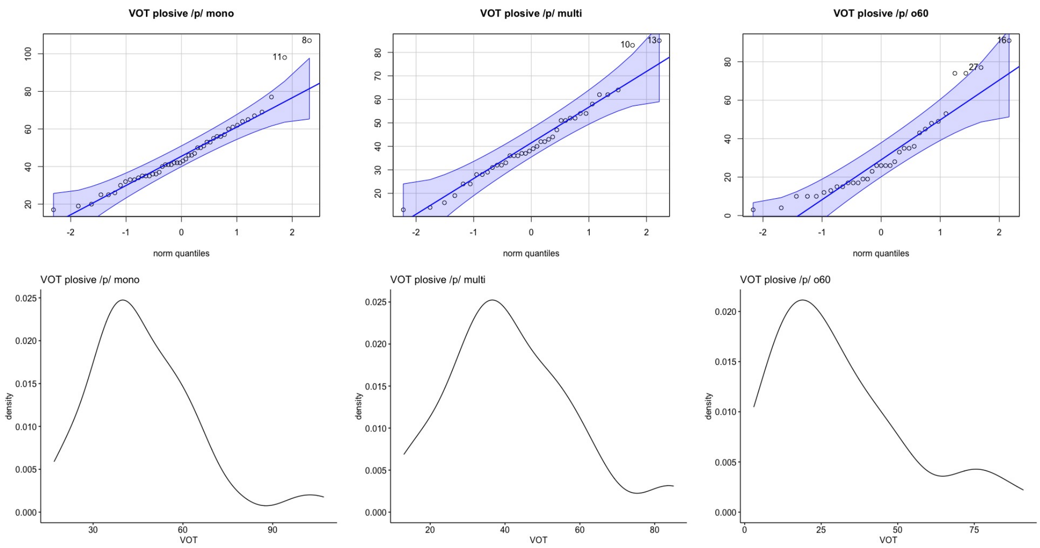


Figure 8.1: Quantile-quantile and density plots for plosive /p/ per group

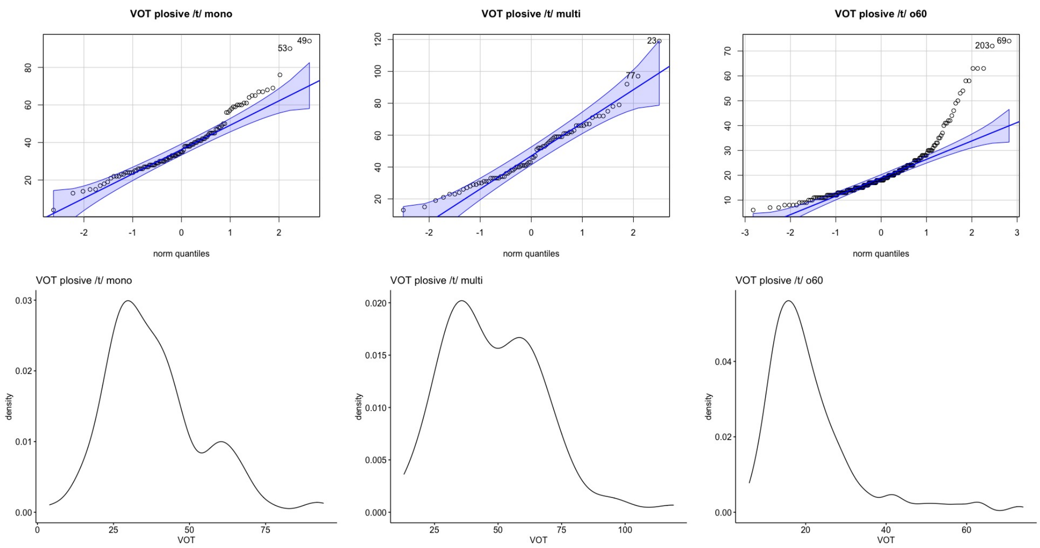


Figure 8.2: Quantile-quantile and density plots for plosive /t/ per group

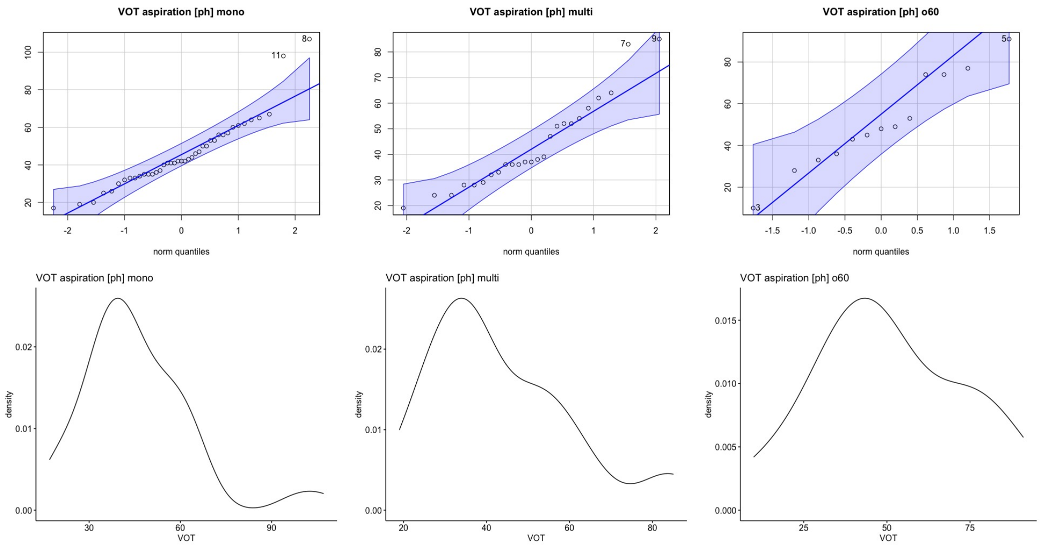


Figure 8.3: Quantile-quantile and density plots for aspiration [p^h] per group

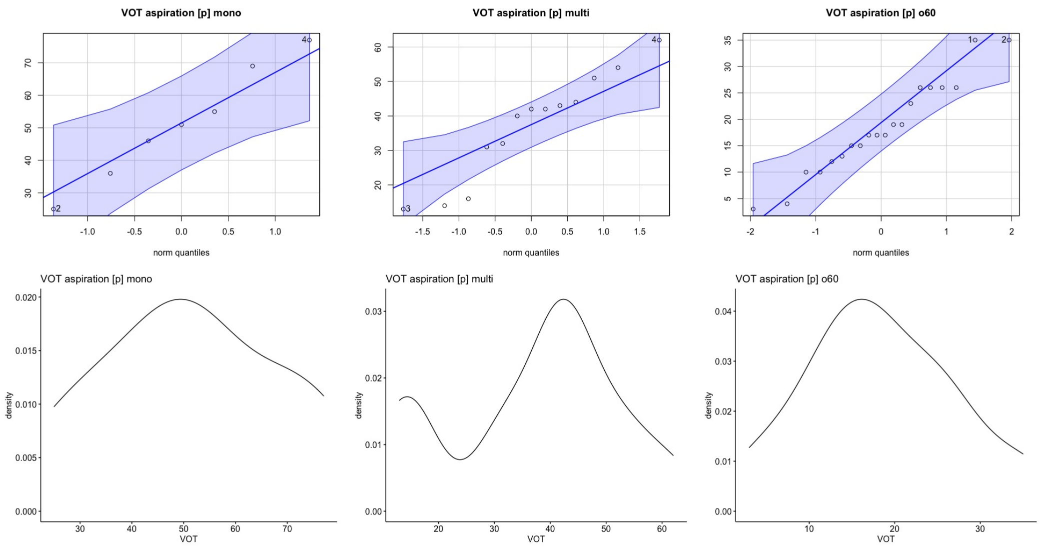


Figure 8.4: Quantile-quantile and density plots for aspiration [p] per group

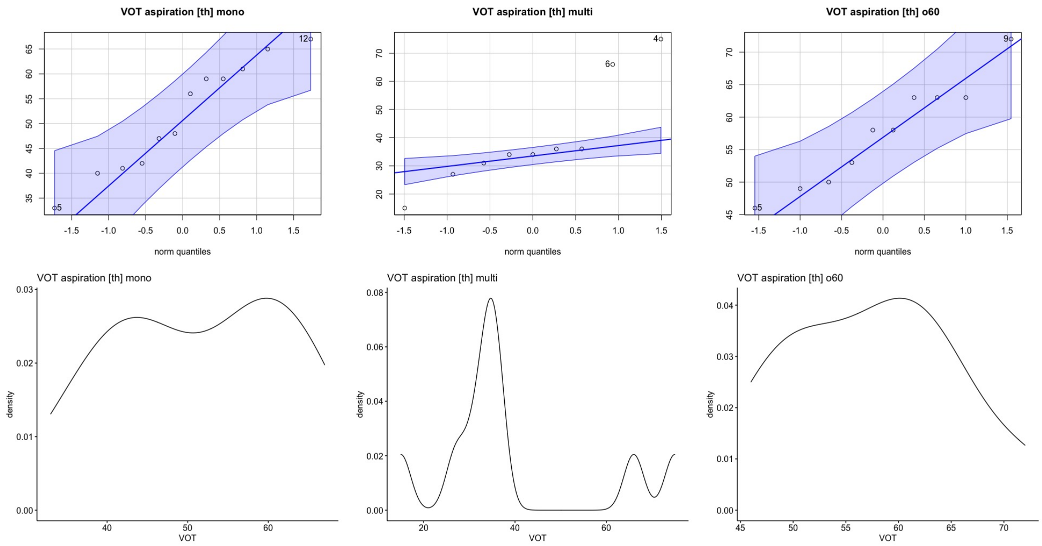


Figure 8.5: Quantile-quantile and density plots for aspiration [tʰ] per group

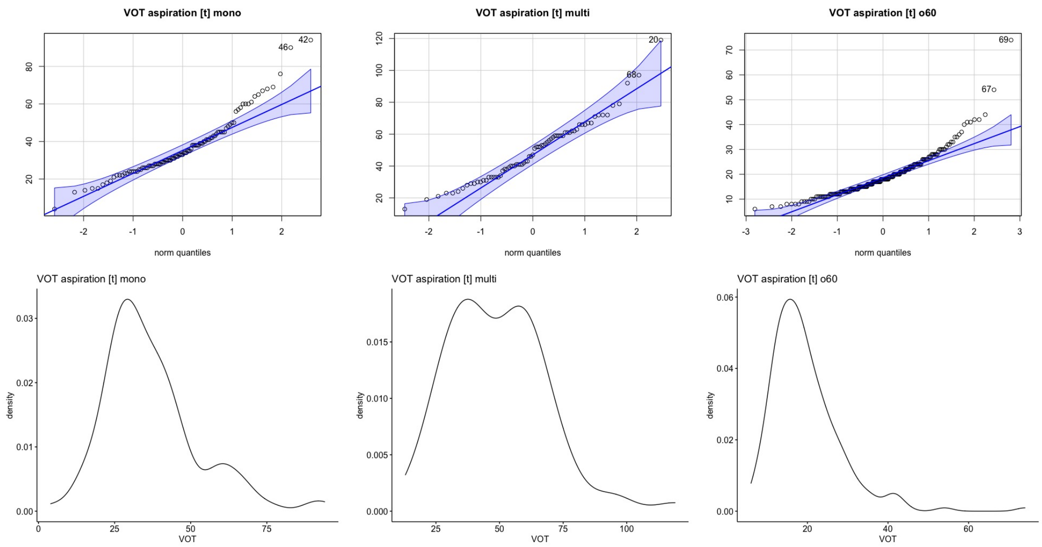


Figure 8.6: Quantile-quantile and density plots for aspiration [t] per group

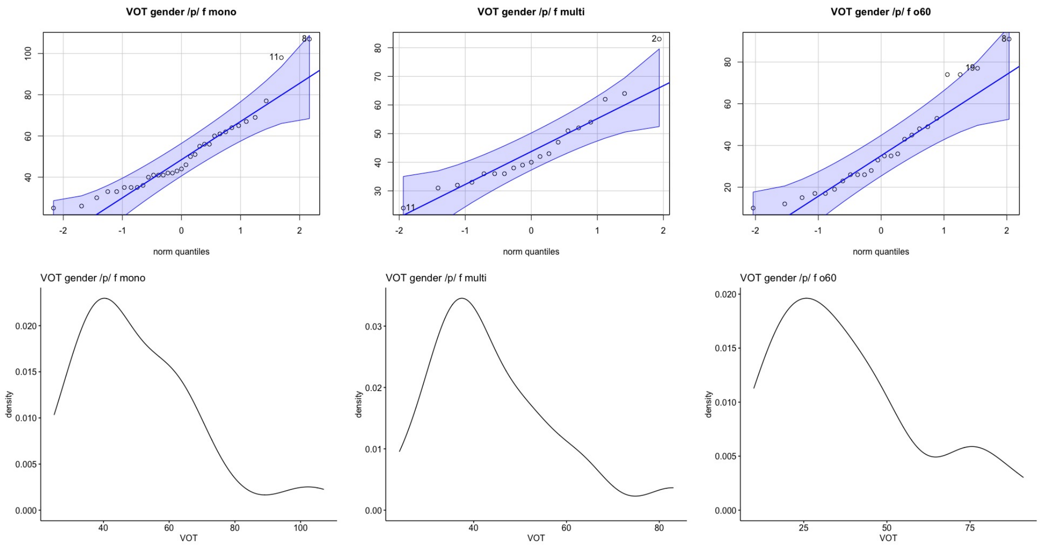


Figure 8.7: Quantile-quantile and density plots for gender /p/ f per group

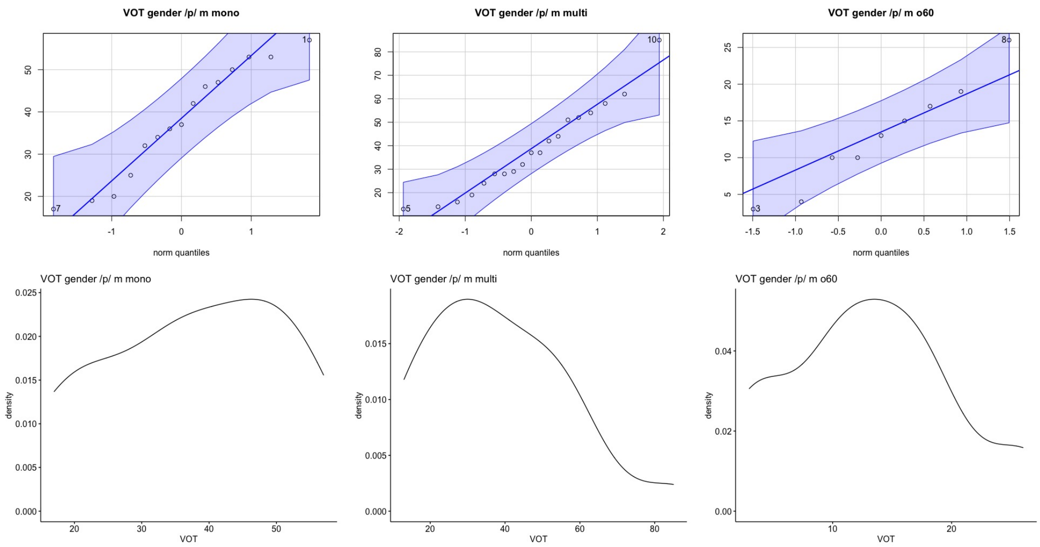


Figure 8.8: Quantile-quantile and density plots for gender /p/ m per group

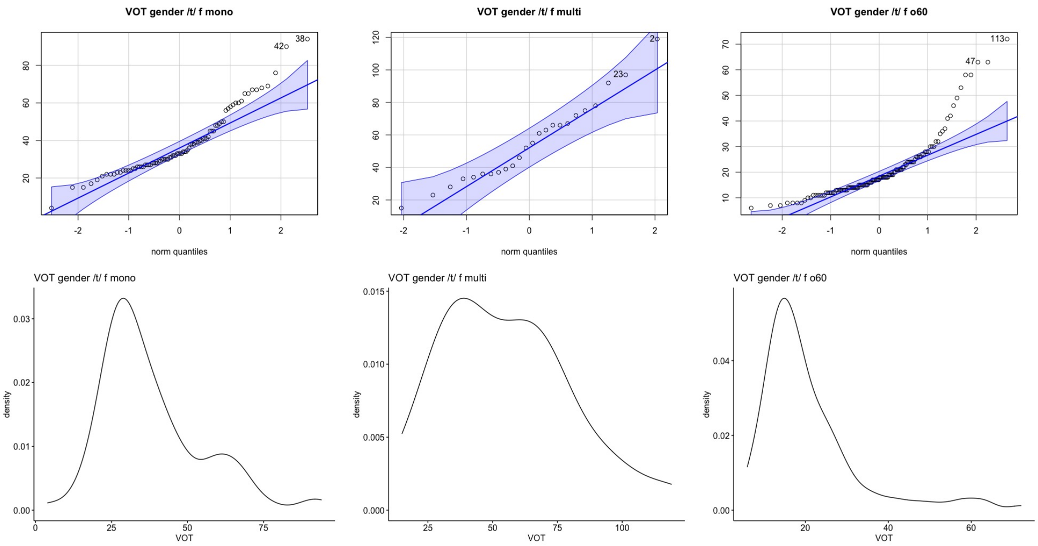


Figure 8.9: Quantile-quantile and density plots for gender /t/ f per group

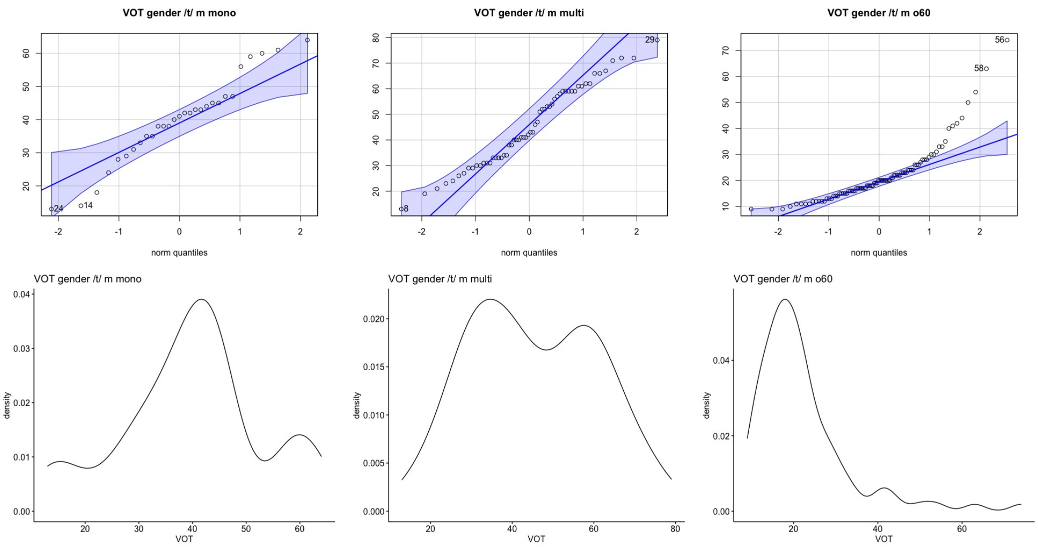


Figure 8.10: Quantile-quantile and density plots for gender /t/ m per group

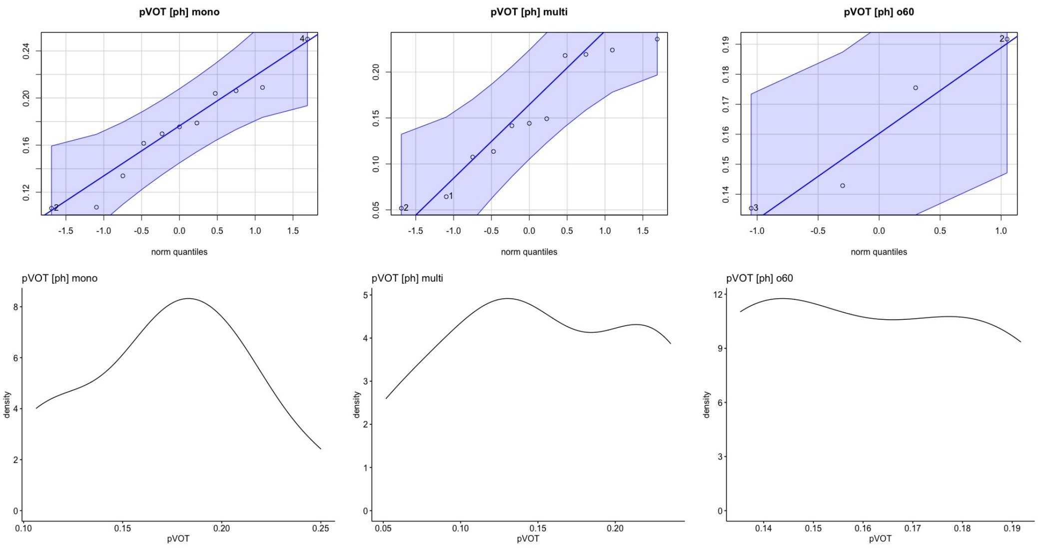


Figure 8.11: Quantile-quantile and density plots for pVOT [p^h] per group

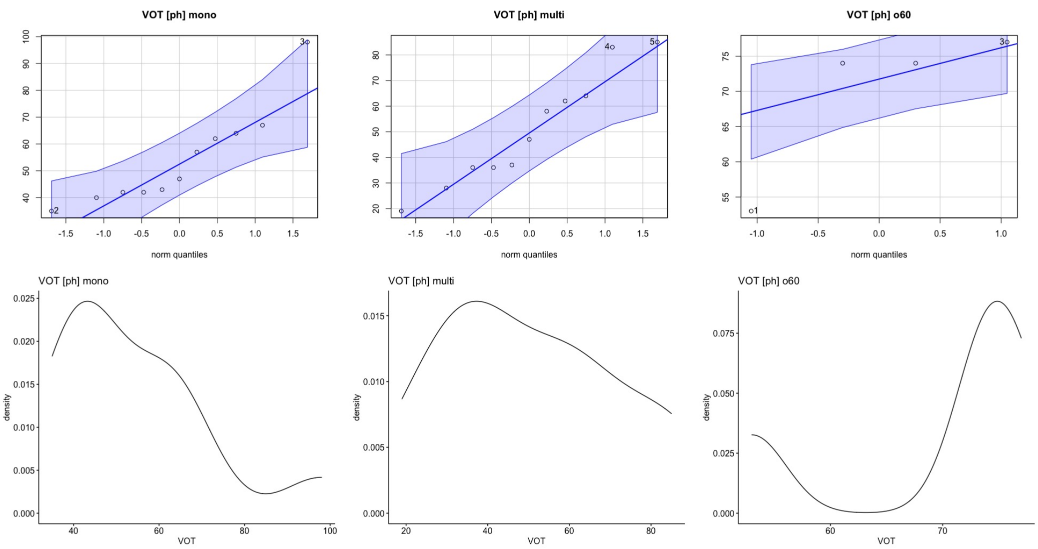


Figure 8.12: Quantile-quantile and density plots for VOT [p^h] per group

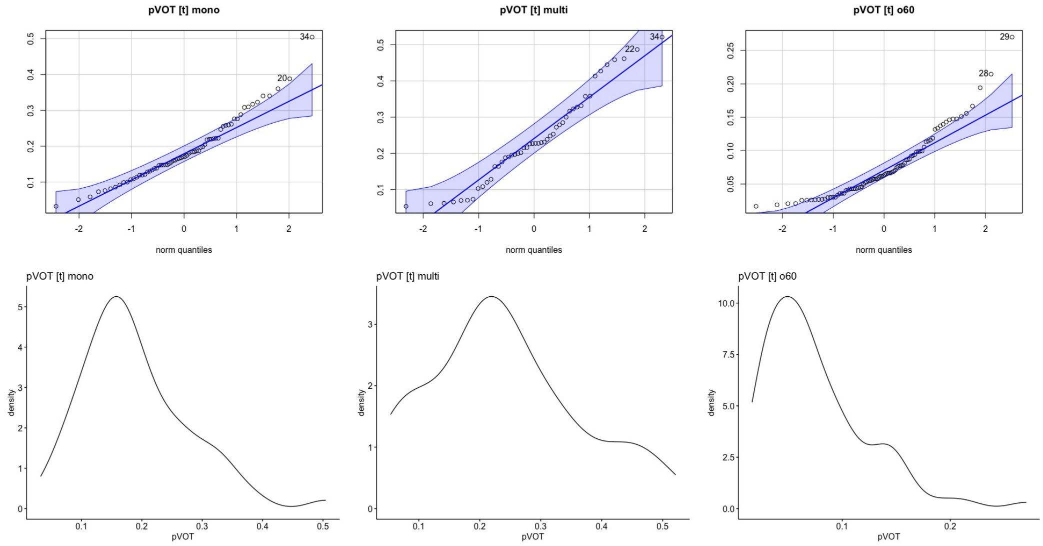


Figure 8.13: Quantile-quantile and density plots for pVOT [t] per group

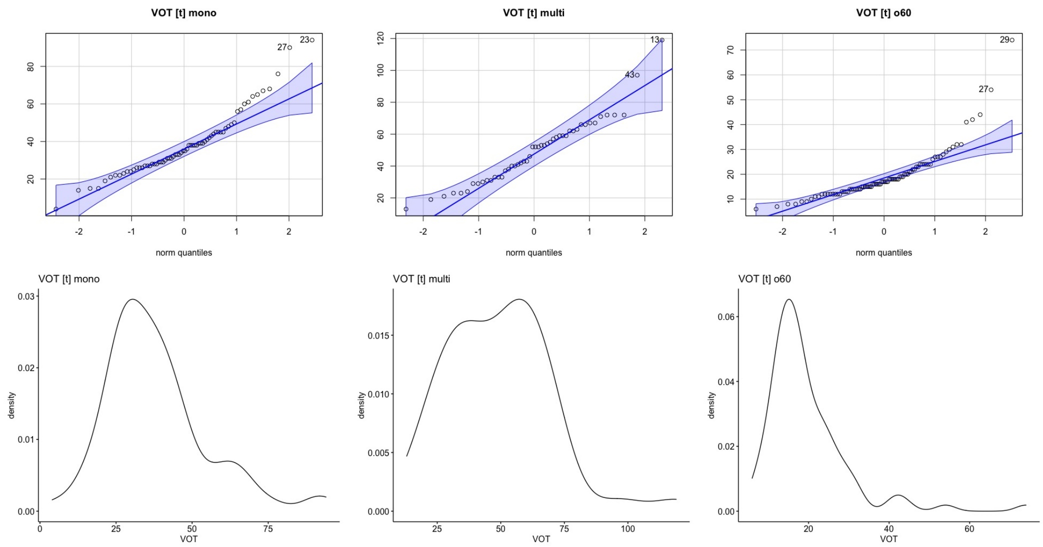


Figure 8.14: Quantile-quantile and density plots for VOT [t] per group

8.2 Boxplots of aspiration comparison within group

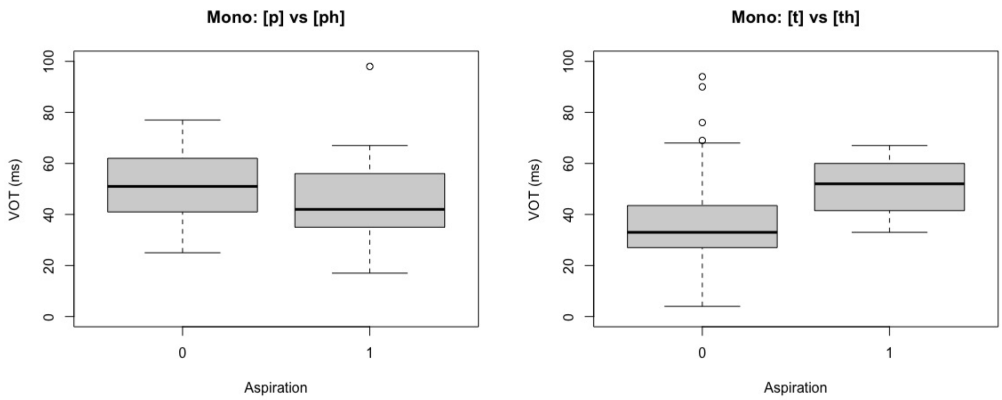


Figure 8.15: Comparison of aspiration per PoA for *mono*

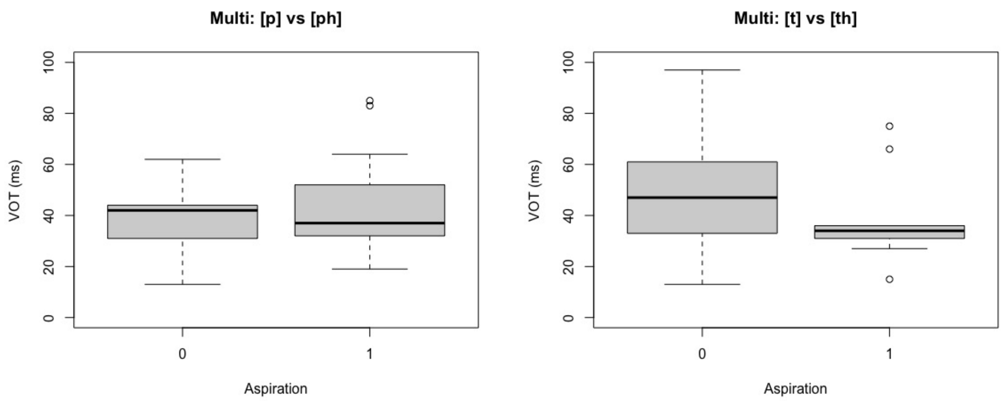


Figure 8.16: Comparison of aspiration per PoA for *multi*

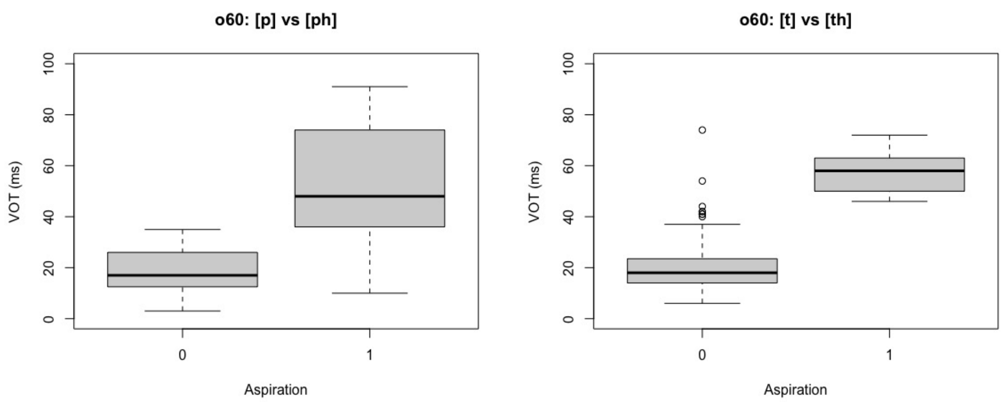


Figure 8.17: Comparison of aspiration per PoA for *o60*

8.3 Boxplots of gender comparison within group

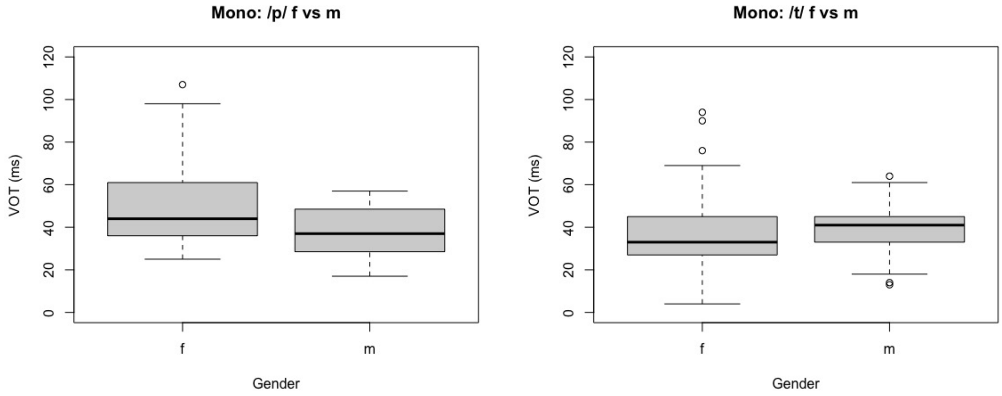


Figure 8.18: Comparison of gender for *mono*

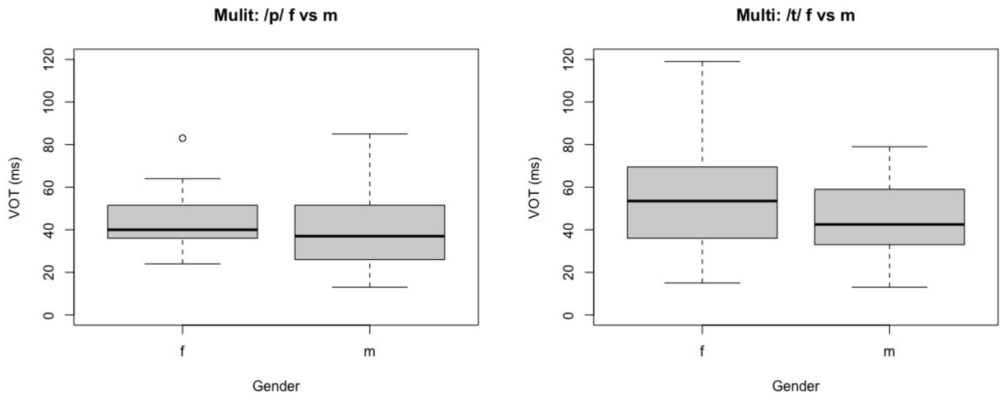


Figure 8.19: Comparison of gender for *multi*

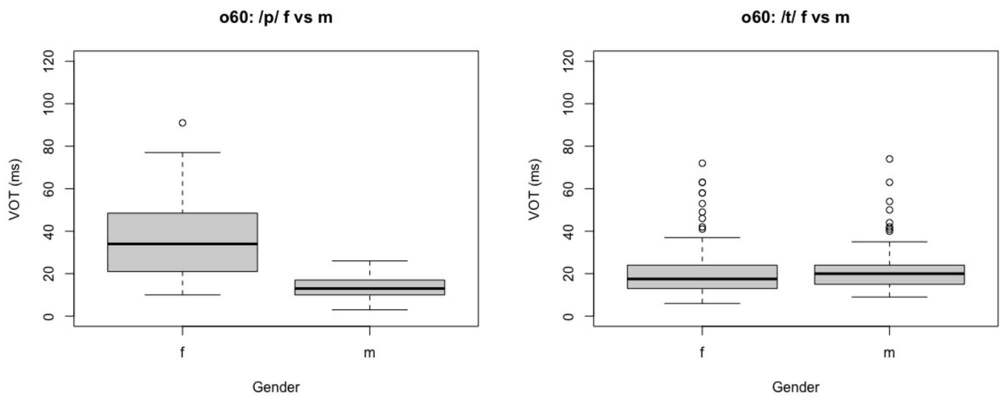
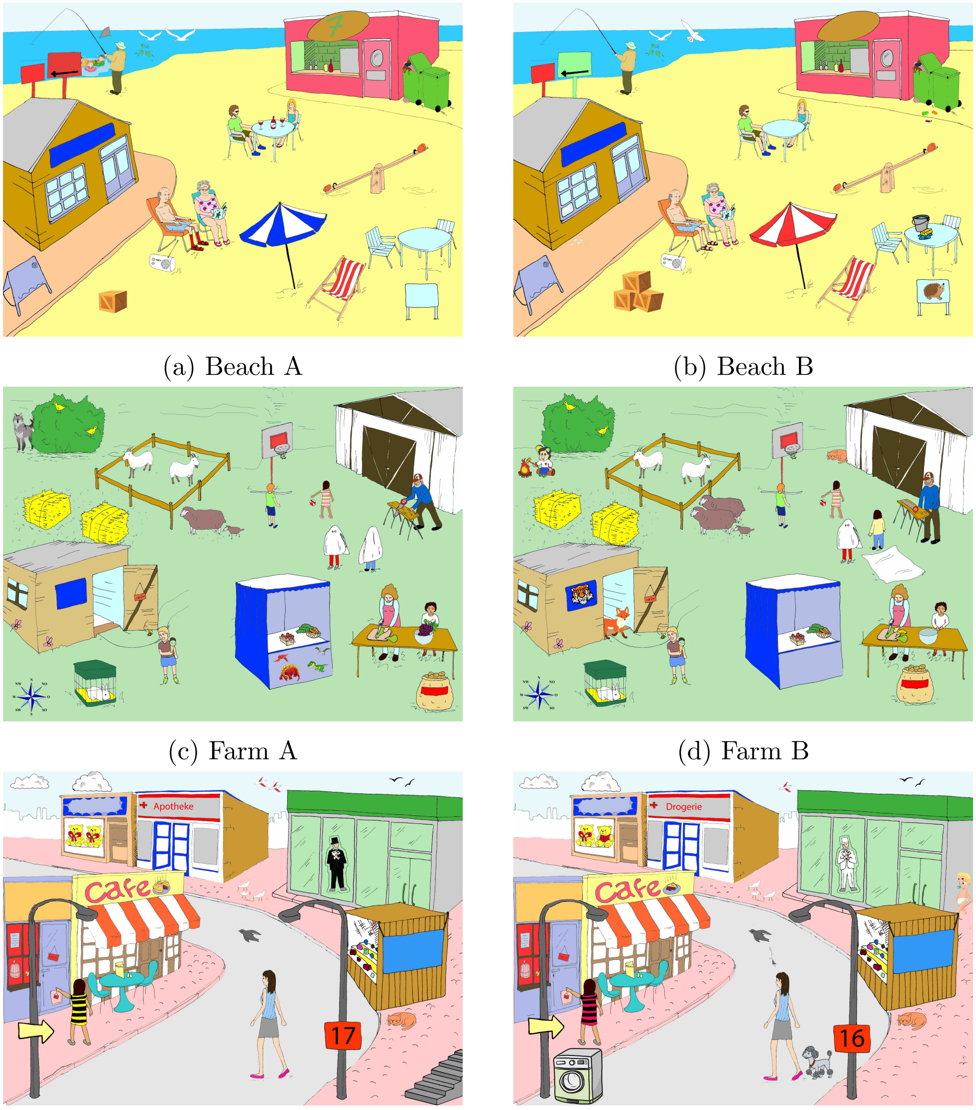


Figure 8.20: Comparison of gender for *o60*

8.4 DiaPix picture pairs



(a) Beach A

(b) Beach B

(c) Farm A

(d) Farm B

(e) Street A

(f) Street B

Figure 8.21: DiaPix used to elicit spontaneous speech by participants

9 Scripts

9.1 Channel and tier extraction from sound files and textgrids

This first Praat script extracts sound channels 1 and 2 from the original sound files of the Diapix corpus and additionally extracts the annotated tiers 1 and 2 from the original textgrid files. The extracted information is renamed and saved in a separate folder.

```
# Script to extract channel 1 and 2 from original sound
# extract annotated tier 1 and 2 from original Textgrids

# clear Objects in Praat
select all
Remove
clearinfo
# define directory of folder containing files and directory of folder to save computed files
dir$ = "/Users/..."
dirToSave$ = "/Users/..."
# Alternative: define directory path manually by user
#dir$ = chooseDirectory$: "Enter a directory"
#dirToSave$ = chooseDirectory$: "Enter a target directory"

# Extract channel 1 and 2 from all wav-sounds
strings = Create Strings as file list: "list", dir$ + "/*.wav"
numberOfSounds = Get number of strings
for ifile to numberOfSounds
   selectObject: strings
   soundName$ = Get string: ifile

   # define variable to current ifile to select current object
   currentSound = Read from file: dir$ + "/" + soundName$

   # Extract first channel = soundwave of first speaker
   Extract one channel: 1

   # rename soundfile and save it to new folder
   newSoundCh1$ = left$ (soundName$, 9) + ".wav"
   #appendInfoLine: soundName$
   #appendInfoLine: newSoundCh1$
   Write to WAV file: dirToSave$ + "/" + newSoundCh1$

   # select current sound to extract soundwave of second speaker and rename after second speaker
   selectObject: currentSound
   Extract one channel: 2
   newSoundCh2$ = left$ (soundName$, 3) + mid$ (soundName$, 10, 6) + ".wav"
   #appendInfoLine: soundName$
   #appendInfoLine: newSoundCh2$
   Write to WAV file: dirToSave$ + "/" + newSoundCh2$
endfor

# Extract annotated tier 1 and 2 from all textgrids
# also rename and save tiers in separate textgrids in target folder
strings = Create Strings as file list: "list", dir$ + "/*.TextGrid"
numberOfTG = Get number of strings

for ifile to numberOfTG
   selectObject: strings
   tgName$ = Get string: ifile
   currentTG = Read from file: dir$ + "/" + tgName$

   # Extract, rename and save first tier as textgrid
   Extract one tier: 1
   #appendInfoLine: tgName$
   newTgTier1$ = left$ (tgName$, 9) + ".TextGrid"
   Save as text file: dirToSave$ + "/" + newTgTier1$

   # Extract, rename and save second tier as Textgrid
   selectObject: currentTG
```

```
    Extract one tier: 2
    #appendInfoLine: tgName$
    newTgTier2$ = left$ (tgName$, 3) + mid$ (tgName$, 10, 6) + ".TextGrid"
    Save as text file: dirToSave$ + "/" + newTgTier2$
endfor
```

9.2 Find and extract plosives from textgrids

The second Praat script searches for bilabial and alveolar fortis plosives /p/ and /t/ followed by a vowel in the annotated tier with regex and outputs a table with plosives, the context (i.e. the position of the plosive within the word), the preceding character of the plosive (white space or placeholder "x" for anything else), the complete annotation of the searched segment and the timecode found plosive.

```
# Script to extract all plosives (t and p) from TextGrids and export information in a table

# clear Objects in Praat
select all
clearinfo
# define directory of folder containing files and directory of folder to save computed files
dir$ = "/Users/..."
# Create table for output and write header of table
outfile$ = "/Users/.../plosivesFromDiapix.txt"
writeFileLine: outfile$, "Speaker", tab$, "Plosive", tab$, "ContextCode", tab$, "Context", tab$,
... "PrecedingPlosive", tab$, "Annotation", tab$, "Timecode"

# Loop through all textgrids in folder to find plosives
strings = Create Strings as file list: "list", dir$ + "/*.TextGrid"
numberOfTG = Get number of strings
tCounter = 0
pCounter = 0
plosiveCounter = 0

for ifile to numberOfTG
    selectObject: strings
    tgName$ = Get string: ifile
    currentTG = Read from file: dir$ + "/" + tgName$

    # extract speakercode (e.g. Ze03f) from filename (DIA_Ze03f.Textgrid)
    speaker$ = mid$ (tgName$, 5, 5)
    appendInfoLine: speaker$

    # count number of intervals in annotated tier (tier 1) in current Textgrid
    noOfIntervals = Get number of intervals: 1

    # loop through all intervals
    for currentInt from 1 to noOfIntervals
        # Get annotation of current interval
        annotation$ = Get label of interval: 1, currentInt

        ### P ###

        # check for 1 or more Pp at start of annotated segment, followed by vowel
        if index_regex(annotation$, "^[Pp]+[aäe èio öuü]")
            match = index_regex(annotation$, "^[Pp]+[aäe èio öuü]")
            appendInfoLine: match

            contextCode$ = "^PV"
            context$ = "Anlaut"
            preceding$ = "White Space"
            appendInfoLine: "p", tab$, contextCode$, tab$, annotation$

            # get timecode of starting time of current interval
            # and round timecode to nearest integer for shorter timestamp without decimals
            timecode = Get start time of interval: 1, currentInt
            timecodeRound = round (timecode)

            # Output: Write information to outfile
            appendFileLine: outfile$, speaker$, tab$, "p", tab$, contextCode$, tab$, context$,
            ... tab$, preceding$, tab$, annotation$, tab$, timecodeRound

            pCounter = pCounter + 1
            plosiveCounter = plosiveCounter + 1
        endif

        # check for 1 or more Pp at word beginning, followed by vowel (\s = whitespace)
        if index_regex(annotation$, "\s[Pp]+[aäe èio öuü]")
```

```
    match = index_regex(annotation$, "\s[Pp]+[aäeèioöuü]")
    appendInfoLine: match
    contextCode$ = "_PV"
    context$ = "Anlaut"
    preceding$ = "White Space"
    appendInfoLine: "p", tab$, contextCode$, tab$, annotation$
    timecode = Get start time of interval: 1, currentInt
    timecodeRound = round (timecode)
    appendFileLine: outfile$, speaker$, tab$, "p", tab$, contextCode$, tab$, context$,
    ... tab$, preceding$, tab$, annotation$, tab$, timecodeRound
    pCounter = pCounter + 1
    plosiveCounter = plosiveCounter + 1
endif

# check for 1 or more Pp, followed by vowel, midword
if index_regex(annotation$, "[Pp]+[aäeèioöuü]") & !index_regex(annotation$,
... "[Pp]+[aäeèioöuü]\s") & !index_regex(annotation$, "\s[Pp]+[aäeèioöuü]")
    match = index_regex(annotation$, "[Pp]+[aäeèioöuü]") & !index_regex(annotation$,
    ... "[Pp]+[aäeèioöuü]\s") & !index_regex(annotation$, "\s[Pp]+[aäeèioöuü]")
    appendInfoLine: match
    contextCode$ = "PV"
    context$ = "Inlaut"
    preceding$ = "x"
    appendInfoLine: "p", tab$, contextCode$, tab$, annotation$
    timecode = Get start time of interval: 1, currentInt
    timecodeRound = round (timecode)
    appendFileLine: outfile$, speaker$, tab$, "p", tab$, contextCode$, tab$, context$,
    ... tab$, preceding$, tab$, annotation$, tab$, timecodeRound
    pCounter = pCounter + 1
    plosiveCounter = plosiveCounter + 1
endif

# check for 1 or more Pp, followed by vowel, wordend (\s = whitespace)
if index_regex(annotation$, "[Pp]+[aäeèioöuü]\s")
    match = index_regex(annotation$, "[Pp]+[aäeèioöuü]\s")
    appendInfoLine: match
    contextCode$ = "PV_"
    context$ = "Auslaut"
    preceding$ = "x"
    appendInfoLine: "p", tab$, contextCode$, tab$, annotation$
    timecode = Get start time of interval: 1, currentInt
    timecodeRound = round (timecode)
    appendFileLine: outfile$, speaker$, tab$, "p", tab$, contextCode$, tab$, context$,
    ... tab$, preceding$, tab$, annotation$, tab$, timecodeRound
    pCounter = pCounter + 1
    plosiveCounter = plosiveCounter + 1
endif

### T ###

# check for 1 or more Tt at beginning of string, followed by vowel (\s = whitespace)
if index_regex(annotation$, "^[Tt]+[aäeèioöuü]")
    match = index_regex(annotation$, "^[Tt]+[aäeèioöuü]")
    appendInfoLine: match
    contextCode$ = "^TV"
    context$ = "Anlaut"
    preceding$ = "White Space"
    appendInfoLine: "t", tab$, contextCode$, tab$, annotation$
    timecode = Get start time of interval: 1, currentInt
    timecodeRound = round (timecode)
    appendFileLine: outfile$, speaker$, tab$, "t", tab$, contextCode$, tab$, context$,
    ... tab$, preceding$, tab$, annotation$, tab$, timecodeRound
    tCounter = tCounter + 1
    plosiveCounter = plosiveCounter + 1
endif

# check for 1 or more Tt at word beginning, followed by vowel (\s = whitespace)
if index_regex(annotation$, "\s[Tt]+[aäeèioöuü]")
    match = index_regex(annotation$, "\s[Tt]+[aäeèioöuü]")
    appendInfoLine: match
    contextCode$ = "_TV"
    context$ = "Anlaut"
    preceding$ = "White Space"
    appendInfoLine: "t", tab$, contextCode$, tab$, annotation$
    timecode = Get start time of interval: 1, currentInt
    timecodeRound = round (timecode)
    appendFileLine: outfile$, speaker$, tab$, "t", tab$, contextCode$, tab$, context$,
    ... tab$, preceding$, tab$, annotation$, tab$, timecodeRound
    tCounter = tCounter + 1
    plosiveCounter = plosiveCounter + 1
endif
```

```
      # check for 1 or more Tt, followed by vowel, midword
      if index_regex(annotation$, "[Tt]+[aäe èioöuü]") & !index_regex(annotation$,
      ... "[Tt]+[aäe èioöuü]\s") & !index_regex(annotation$, "\s[Tt]+[aäe èioöuü]")
         match = index_regex(annotation$, "[Tt]+[aäe èioöuü]") & !index_regex(annotation$,
         ... "[Tt]+[aäe èioöuü]\s") & !index_regex(annotation$, "\s[Tt]+[aäe èioöuü]")
         appendInfoLine: match
         contextCode$ = "TV"
         context$ = "Inlaut"
         preceding$ = "x"
         appendInfoLine: "t", tab$, contextCode$, tab$, annotation$
         timecode = Get start time of interval: 1, currentInt
         timecodeRound = round (timecode)
         appendFileLine: outfile$, speaker$, tab$, "t", tab$, contextCode$, tab$, context$,
         ... tab$, preceding$, tab$, annotation$, tab$, timecodeRound
         tCounter = tCounter + 1
         plosiveCounter = plosiveCounter + 1
      endif

      # check for 1 or more Tt, followed by vowel, wordend (\s = whitespace)
      if index_regex(annotation$, "[Tt]+[aäe èioöuü]\s")
         match = index_regex(annotation$, "[Tt]+[aäe èioöuü]\s")
         appendInfoLine: match
         contextCode$ = "TV_"
         context$ = "Auslaut"
         preceding$ = "x"
         appendInfoLine: "t", tab$, contextCode$, tab$, annotation$
         timecode = Get start time of interval: 1, currentInt
         timecodeRound = round (timecode)
         appendFileLine: outfile$, speaker$, tab$, "t", tab$, contextCode$, tab$, context$,
         ... tab$, preceding$, tab$, annotation$, tab$, timecodeRound
         tCounter = tCounter + 1
         plosiveCounter = plosiveCounter + 1
      endif
   endfor
endfor
```

9.3 Find and extract plosives from Praat output

```
#!/usr/bin/python
# -*- coding: utf-8 -*-
# Terminal : python 3 findPlosivesInLine.py
# cwatte

# Script takes table from extractPlosives.praat and gather following information:
# PlosiveCounter, PrecedingWord, PrecedingLetter, Plosive, FollowingVowel,
# WordWithInitialPlosive    , AnnotationComplete
# output in outfile : 1 din    n  T  e  Teddibäär   hätt din Teddibäär so Schläiffe

import re, codecs, os, sys

outfile = open('FindPlosivesFromPraatOutput.txt', 'w')
outfileDict = open('Plosives.txt', 'w')
plosiveCounter = 0
plosiveDict = {}
with codecs.open('TestU8.txt', 'r', encoding='utf-8') as infile:
   for line in infile:
      # convert line to list
      words = line.split()
      # w = individual words in wordlist
      precedingW = 'Plosive is initial in annotation'

      for w in words:
         # check if first letter is a plosive
         if (w[0] == 'P') or (w[0] == 'p') or (w[0] == 'T') or (w[0] == 't'):
            # check if plosive is in plosiveDict: add it, or add counter
            if w not in plosiveDict.keys():
               plosiveDict[w] = 1
            else:
               plosiveDict[w] += 1

            # collect preceding word and letter
            if precedingW == 'Plosive is initial in annotation':
               precedingLetter = 'unknown'
            else:
               precedingLetter = precedingW[-1]
```

```
            # collect vowel following plosive; catch error if plosive is standalone letter
            if len(w) == 1:
                followingVowel = "white space"
            else:
                followingVowel = w[1]

            # display collected info to terminal and outfile
            # print (precedingW, '\t', precedingLetter, '\t', w[0], '\t', followingVowel, '\t', w)

            plosiveCounter += 1
            counterstr = str(plosiveCounter)
            outfile.write(counterstr)
            outfile.write('\t')
            outfile.write(precedingW)
            outfile.write('\t')
            outfile.write(precedingLetter)
            outfile.write('\t')
            outfile.write(w[0])
            outfile.write('\t')
            outfile.write(followingVowel)
            outfile.write('\t')
            outfile.write(w)
            outfile.write('\t')
            outfile.write(line)

        # save current word that will potentially be preceding plosive
        precedingW = w
# write counted plosives in outfile "outfileDict"
sortedPlosives = sorted(plosiveDict, key=plosiveDict.get, reverse=True)
for w in sortedPlosives:
    plosivecount = str(plosiveDict[w])
    print (w, '\t', plosivecount, '\n')
    outfileDict.write(w)
    outfileDict.write('\t')
    outfileDict.write(plosivecount)
    outfileDict.write('\n')
infile.close()
outfile.close()
outfileDict.close()
```

9.4 Extract annotation and sound files for Webmaus

```
# Script to searches for segments containing specific plosives in all sound files
# extracts sound of matched segments and stores it in new renamed wav file
# extracts annotation of matched segments and stores it in new renamed txt file
# additionally export information in a table (plosive counter, name of sound file)
# to control if all sound and annotation file have been processed

# clear Objects in Praat
select all
nocheck Remove
clearinfo
# define directory of folder containing files and directory of folder to save computed files
dir$ = "/Users/..."
outdir$ = "/Users/..."
# Create table and file for output
outfile$ = "/Users/.../_ExtractedSoundNames.txt"
outtable$ = "/Users/.../_TableExtractedSoundNames.txt"

# pop up window to enter plosive search
form Plosive Search
     comment Enter plosive to be searched.
     word Plosive plosive
endform
appendInfoLine: plosive$
pCounter = 0

appendFileLine: outfile$, ""
appendFileLine: outfile$, plosive$
appendFileLine: outfile$, ""

# Create list of Speakerfiles in directory
strings = Create Strings as file list: "list", dir$ + "/*.wav"
number_of_files = Get number of strings

# Loop through all files in directory by selecting and reading wav file as well as textgrid
for file to number_of_files

    # Select and read wav-file in directory
```

```
    selectObject: strings
    speakerfile$ = Get string: file
    sound = Read from file: dir$ + "/" + speakerfile$
    speaker$ = speakerfile$ - ".wav"
    appendInfoLine: "speakerfile", tab$, speakerfile$

    # Read textgrid of selected sound in directory
    textgrid = Read from file: dir$ + "/" + speakerfile$ - "wav" + "TextGrid"

    # select sound and textgrid pair
    selectObject: sound, textgrid

    # cut out sound with plosive in annotation
    nowarn Extract intervals where: 1, "yes", "contains a word equal to", plosive$
    #appendInfoLine: plosive$

    # store found sounds
    selectedIntervals = numberOfSelected ("Sound")
    for i to selectedIntervals
       intervals[i] = selected("Sound", i)
    endfor

    # loop through stored found sound
    for i to selectedIntervals
       selectObject: intervals[i]

       # select sounds individually and rename (delete "DIA_" in front for better overview)
       # and save sound to output directory
       name$ = selected$ ("Sound")
       length = length (name$)

       # only select found plosives which have new longer names in Object window
       if length > 10
          idname$ = replace$ (name$, "DIA_", "", 0)
          Write to WAV file: outdir$ + "/" + idname$ + ".wav"

          # get start and end timecode of sound to extract annotation in txt
          start = Get start time
          end = Get end time
          #appendInfoLine: start
          #appendInfoLine: end

          # find annotation with timecodes and store it in txt file for Webmaus annotation
          tg$ = left$ (name$, 9)
          appendInfoLine: tg$
          select TextGrid 'tg$'
          Extract part: start, end, "yes"
          annotation$ = Get label of interval: 1, 1
          appendInfoLine: annotation$
          txt$ = outdir$ + "/" + idname$ + ".txt"
          appendFileLine: txt$, annotation$

          # collect info in table for tracking
          pCounter = pCounter + 1
          appendInfoLine: pCounter, tab$, idname$, tab$, plosive$
          appendFileLine: outfile$, pCounter, tab$, idname$, tab$, plosive$, tab$, annotation$
          appendFileLine: outtable$, pCounter, tab$, idname$, tab$, plosive$, tab$, annotation$
       endif
    endfor
endfor
```

9.5 Alter tiers in textgrids

9.5.1 Alter tier of Webmaus output

```
# Script takes textgrid from Webmaus and alters tiers and information in tiers

# clear Objects in Praat
select all
nocheck Remove
clearinfo
# define direcotry of folder containing files and directory of folder to save computed files
dir$ = "/Users/.../inputTG"
outdir$ = "/Users/.../outputTG"
pCounter = 0

# Create list of files in directory to loop through files
```

```
strings = Create Strings as file list: "list", dir$ + "/*.TextGrid"
number_of_files = Get number of strings

for file to number_of_files
   # Select and read textgrid-file in directory
   selectObject: strings
   name$ = Get string: file
   Read from file: dir$ + "/" + name$

   # remove irrelevant tiers
   Remove tier: 5
   Remove tier: 3
   Remove tier: 2

   # add interval and rename existing tiers
   Set tier name: 1, "words"
   Set tier name: 2, "segments"

   # get number of intervals to collect whole utterance in string "sentence$"
   nr_of_words = Get number of intervals: 1
   sentence$ = ""

   for i to nr_of_words
      word$ = Get label of interval: 1, i
      if word$ = ""
         appendInfoLine: "empty"
      else
         sentence$ = sentence$ + " " + word$
      endif
   endfor
   appendInfoLine: sentence$

   # insert new empty tier at top
   Insert interval tier: 1, "sentence"
   # insert text in tier 1 and interval 1:
   Set interval text: 1, 1, sentence$
   Save as text file: outdir$ + "/" + name$
endfor
```

9.5.2 Inserting tier

```
# Script takes textgrid from Webmaus and inserts one tiers

# clear Objects in Praat
select all
nocheck Remove
clearinfo
# define direcotry of folder containing files and directory of folder to save computed files
dir$ = "/Users/.../input_tier"
outdir$ = "/Users/.../output_tier"
pCounter = 0
# Create list of files in directory to loop through files
strings = Create Strings as file list: "list", dir$ + "/*.TextGrid"
number_of_files = Get number of strings

for file to number_of_files
   # Select and read textgrid-file in directory
   selectObject: strings
   name$ = Get string: file
   Read from file: dir$ + "/" + name$

   # insert new empty tier at top
   Insert interval tier: 4, "phases"
   Save as text file: outdir$ + "/" + name$
endfor
```

9.5.3 Duplicate segment tier

```
# duplicate segment tier for single annotation of target word

# clear Objects in Praat
select all
nocheck Remove
clearinfo
# define direcotry of folder containing files and directory of folder to save computed files
dir$ = "/Users/..."
outdir$ = "/Users/..."
pCounter = 0
```

```
# Create list of files in directory to loop through files
strings = Create Strings as file list: "list", dir$ + "/*.TextGrid"
number_of_files = Get number of strings
for file to number_of_files
   # Select and read textgrid-file in directory
   selectObject: strings
   name$ = Get string: file
   Read from file: dir$ + "/" + name$
   # duplicate segment tier and add duplicate right underneath
   Duplicate tier: 3, 4, "targetWord"

   Save as text file: outdir$ + "/" + name$
endfor
```

9.5.4 Insert dummy annotation

```
# Script insert dummy |C|R| annotation and boundaries in phases tier
# to facilitate manual annotation

# clear Objects in Praat
select all
nocheck Remove
clearinfo
# define directory of folder containing files and directory of folder to save computed files
dir$ = "/Users/..."
outdir$ = "/Users/..."
# Create list of files in directory to loop through files
strings = Create Strings as file list: "list", dir$ + "/*.TextGrid"
number_of_files = Get number of strings

for file to number_of_files
   # Select and read textgrid-file in directory
   selectObject: strings
   tg_name$ = Get string: file
   textgrid = Read from file: dir$ + "/" + tg_name$
   selectObject: textgrid

   #get number of intervals in words tier to loop through it
   nr_of_words = Get number of intervals: 2

   # define variables with dummy values
   c_start = 0.1
   burst = 0.2
   r_end = 0.3

   for w to nr_of_words
      # get word
      annot$ = Get label of interval: 2, w

      if index_regex (annot$, "^[PpTt]+[aäe èioöuü]")

         # check if plosive word is in first interval
         # if yes, there is no boundary and dummy c_start is set to 0.11
         if w = 1
            c_start = 0.11
         else
            c_start = Get start time of interval: 2, w
         endif

         appendInfoLine: "C_start= ", c_start

         # in tier targetWords (4) search for plosive interval (same start time as word)
         # and get end boundary of plosive (= r_end)
         plosive_interval = Get interval boundary from time: 4, c_start

         # catch erorr: set interval to 1
         if plosive_interval = 0
            plosive_interval = 1
         endif

         # get end of release time at annotation of plosive in tier 4
         # set dummy burst in the middle of C_start and R_end boundary
         r_end = Get end time of interval: 4, plosive_interval
         annot_plosive$ = Get label of interval: 4, plosive_interval
         burst = r_end - ((r_end-c_start)/2)

         appendInfoLine: plosive_interval
         appendInfoLine: "r_end= ", r_end
         appendInfoLine: "burst= ", burst
         appendInfoLine: "plosive= ", annot_plosive$
```

```
         appendInfoLine: ""

         # insert | C | R | in phases tier
         # set interval text: tier, interval, annot
         Insert boundary: 5, c_start
         Insert boundary: 5, burst
         Insert boundary: 5, r_end
         Set interval text: 5, 2, "C"
         Set interval text: 5, 3, "R"
      endif
   endfor
   Save as text file: outdir$ + "/" + tg_name$
endfor
```

9.6 Cut sound and textgrid at plosive with context

9.6.1 Cut all files automatically

```
# script to cut out plosive with context from complete utterance
# generates new wave file and textgrid: 0.1s + precedingWord + plosiveWord + 0.1s

# clear Objects in Praat
select all
nocheck Remove
clearinfo
# define directory of folder containing files and directory of folder to save computed files
dir$ = "/Users/.../input"
outdir$ = "/Users/.../output"
# Create list of files in directory to loop through files
strings = Create Strings as file list: "list", dir$ + "/*.TextGrid"
number_of_files = Get number of strings

for file to number_of_files
   # Select and read textgrid-file in directory
   selectObject: strings
   tg_name$ = Get string: file
   appendInfoLine: ""
   appendInfoLine: tg_name$

   textgrid = Read from file: dir$ + "/" + tg_name$
   sound = Read from file: dir$ + "/" + tg_name$ - "TextGrid" + "wav"
   selectObject: textgrid

   # loop through intervals in tier "words" to find plosive
   nr_of_words = Get number of intervals: 2
   previous_word$ = ""
   previous_start = 0

   for i to nr_of_words
      # collect text and start time of current interval
      current_word$ = Get label of interval: 2, i
      current_start = Get start time of interval: 2, i
      #appendInfoLine: current_start

      # find interval with plosive
      if index_regex (current_word$, "^[PpTt]+[aäeèioöuü]")
         appendInfoLine: previous_word$, " ", current_word$

         # check if plosive is first interval
         # if yes, collect start time of plosive
         # if no, collect start time of word before plosive
         if previous_word$ = ""
            start = Get start time of interval: 2, i
            appendInfoLine: "Plosive is first"
            appendInfoLine: "starttime = ", start
         else
            start = previous_start
            appendInfoLine: "word before plosive = ", previous_word$
            appendInfoLine: "starttime = ", start
         endif

         # collect endtime of plosive
         plosive_end = Get end time of interval: 2, i
      endif

      previous_word$ = current_word$
      previous_start = current_start
```

```
    endfor

    # cut TG and sound with a little context of 0.1 seconds
    cut_start = start - 0.1
    cut_end = plosive_end + 0.1

    out_tg = Extract part: cut_start, cut_end, "no"
    selectObject: out_tg
    Save as text file: outdir$ + "/" + tg_name$

    selectObject: sound
    out_sound = Extract part: cut_start, cut_end, "rectangular", 1, "no"
    selectObject: out_sound
    Write to WAV file: outdir$ + "/" + tg_name$ - "TextGrid" + "wav"
endfor
```

9.6.2 Save sound and textgrid of manual selection

```
# Script opens all sounds + textgrids in folder and opens them in pairs one at a time
# save new sound and textgrid of manual selection in new folder

# clear Objects in Praat
select all
nocheck Remove
clearinfo
# define direcotry of folder containing files and directory of folder to save computed files
indir$ = "/Users/..."
outdir$ = "/Users/..."
strings = Create Strings as file list: "list", indir$ + "/*.TextGrid"
number_of_files = Get number of strings

for file to number_of_files
    selectObject: strings
    tg_name$ = Get string: file

    # read and select current sound + textgrid and open them in pairs
    textgrid = Read from file: indir$ + "/" + tg_name$
    sound = Read from file: indir$ + "/" + tg_name$ - "TextGrid" + "wav"
    selectObject: sound, textgrid

    Edit
    pause

    # CHOOSE SELECTION AND DO:
    # File / Extract selected sound (time from 0)
    # File / Extract selected TextGrid (time from 0)
    # TO CREATE NEW OBJECTS IN PRAAT

    # save newly created textgrid and sound
    newtg = selected (-1)
    selectObject: newtg
    Save as text file: outdir$ + "/" + tg_name$
    selectObject: "Sound untitled"
    Write to WAV file: outdir$ + "/" + tg_name$ - "TextGrid" + "wav"

    # delete both processed files and newly extracted parts
    # in Praat objects window except string list
    select all
        minus Strings list
        Remove
endfor
```

9.7 Measurement calculation

9.7.1 Move boundaries to nearest zero crossing

```
# Script set all boundaries of targetWord and phases tier to nearest zero crossing
# boundaries of onset and offset of plosive in phases tier are added to targetWord tier

select all
nocheck Remove
clearinfo
# define direcotry of folder containing files and directory of folder to save computed files
dir$ = "/Users/..."
outdir$ = "/Users/..."
```

```
strings = Create Strings as file list: "list", dir$ + "/*.TextGrid"
number_of_files = Get number of strings

for file to number_of_files
   selectObject: strings
   tg_name$ = Get string: file
   textgrid = Read from file: dir$ + "/" + tg_name$
   sound = Read from file: dir$ + "/" + tg_name$ - "TextGrid" + "wav"

   # select textgrid and insert new intervals for adjusted boundaries
   selectObject: textgrid
   zeroWord = Insert interval tier: 5, "targetWordZero"
   zeroPhases = Insert interval tier: 7, "phasesZero"

   # loop through tier "phases" and store time codes of boundaries |C|R| in phases tier
   currentTier = 6
   p_start = Get end point: currentTier, 1
   burst = Get end point: currentTier, 2
   p_end = Get end point: currentTier, 3

   # select sound and find nearest zero crossing of |C|R| boundaries
   select sound
   p_start_nearZero = Get nearest zero crossing: 1, p_start
   burst_nearZero = Get nearest zero crossing: 1, burst
   p_end_nearZero = Get nearest zero crossing: 1, p_end

   # go back to TG and insert new boundaries and annotation
   # in new_phasesTier and new_targetWordTier
   # also insert plosive annotation
   select textgrid
   Insert boundary: 7, p_start_nearZero
   Insert boundary: 7, burst_nearZero
   Insert boundary: 7, p_end_nearZero
   Set interval text: 7, 2, "C"
   Set interval text: 7, 3, "R"

   Insert boundary: 5, p_start_nearZero
   Insert boundary: 5, p_end_nearZero
   plosive_annot$ = Get label of interval: 4, 2
   Set interval text: 5, 2, plosive_annot$

   # loop through remaining annotations in tier "targetWord"
   # then insert adjusted boundaries + annotation in new tier
   currentTier = 4
   nr_of_boundaries = Get number of intervals: currentTier

   # nr-1 because always end point of intervals are processed
   for b from 3 to nr_of_boundaries - 1
      # get annotation of current interval
      annot$ = Get label of interval: currentTier, b

      # get end timecode of boundary of current intervall
      right_boundary = Get end point: currentTier,  b

      # go to sound file and get nearest zero crossing of current boundary
      select sound
      timeNearestZero = Get nearest zero crossing: 1, right_boundary

      # go back to textgrid to enter new boundary at zero crossing and annotation
      # in newly created tier underneath
      select textgrid
      Insert boundary: 5, timeNearestZero
      Set interval text: 5, b, annot$
   endfor

   # remove tiers with non zero crossing boundaries
   Remove tier: 6
   Remove tier: 4
   Save as text file: outdir$ + "/" + tg_name$
endfor
```

9.7.2 Durational measurements

```
# Script calculates durations of closure and release phase and normalized VOT

# clear Objects in Praat
select all
nocheck Remove
clearinfo
```

```
# define direcotry of folder containing files and directory of folder to save computed files
dir$ = "/Users/..."

# Create table for output
outtable$ = "/Users/.../_TableDuration.txt"
writeFileLine: outtable$, "NR",tab$,"SPEAKER",tab$,"ID",tab$, "Annotation",tab$, "Closure (s)",tab$,"Closure (ms)",
tab$, "Release (s)",tab$,"Release (ms)",tab$, "pVOT",tab$, "Total Plosive (s)",tab$,"Total Plosive (ms)", tab$,
"TargetW - P (s)", tab$, "TargetW - P (ms)"

# Create list of TextGrids in directory
strings = Create Strings as file list: "list", dir$ + "/*.TextGrid"
number_of_files = Get number of strings

for file to number_of_files
   selectObject: strings
   tg_name$ = Get string: file

   # collect metadata from filename
   nr$ = left$(tg_name$, 3)
   speaker$ = mid$(tg_name$, 5, 5)
   length = length(tg_name$ - ".TextGrid")
   id_plosive$ = mid$(tg_name$, 11, length) - ".TextGrid"

   textgrid = Read from file: dir$ + "/" + tg_name$

   #####
   # calculate C,R and total plosive duration
   # convert from s to ms and cut off decimal digits
   #####
   closure = Get total duration of intervals where: 5, "is equal to", "C"
   release = Get total duration of intervals where: 5, "is equal to", "R"

   c$ = fixed$(closure, 5)
   clos = number(c$)
   closure = round(clos * 1000)

   r$ = fixed$(release, 5)
   rel = number(r$)
   release = round(rel * 1000)

   total_plosive = clos + rel
   total_plosive_round = closure + release

   #####
   # calculate VOT normalization
   # = duration of VOT (release) / duration of target word without plosive
   #####

   p_annot$ = Get label of interval: 4, 2
   p_end = Get end time of interval: 4, 2
   nr_intervals_targetWord = Get number of intervals: 4
   last_boundary = Get start time of interval: 4, nr_intervals_targetWord

   dur_targetWord_MinusPlosive = last_boundary - p_end
   dur$ = fixed$(dur_targetWord_MinusPlosive, 5)
   dur = number(dur$)
   dur_minusPlosive = round(dur * 1000)

   pVOT = release / dur_minusPlosive

   # enter all collected information in output table
   appendFileLine: outtable$, nr$, tab$, speaker$, tab$, id_plosive$, tab$, p_annot$,tab$, c$,tab$,closure,tab$,
    r$,tab$,release,tab$, pVOT,tab$,total_plosive,tab$,total_plosive_round,tab$, dur$,tab$,dur_minusPlosive
endfor
```

9.8 Other utility scripts

9.8.1 Sort and count plosives

```
#!/usr/bin/python
# -*- coding: utf-8 -*-
# Terminal : python 3 findPlosivesInLine.py
# cwatte
# Script takes "WordWithInitialPlosive"-column from DIAPIX_ALL tableInsertModeElements
# collects all plosives in dictionary and counts their occurrence
# outputs a table with all plosives and occurrence count, sorted after highest occurrence
```

```
import re, codecs, os, sys
outfileDict = open('CountdownTokenPlosive.txt', 'w')
plosiveDict = {}
with codecs.open('WordwithInitialPlosives.txt', 'r', encoding='utf-8') as infile:
   for plosive in infile:
      # remove \n from string and make case insensitive
      plosive = plosive.lower().rstrip()
      if plosive not in plosiveDict.keys():
         plosiveDict[plosive] = 1
      else:
         plosiveDict[plosive] += 1

# write counted plosives in outfile "outfileDict"
sortedPlosives = sorted(plosiveDict, key=plosiveDict.get, reverse=True)
counter = 0
for w in sortedPlosives:
   counter += plosiveDict[w]
   plosivecount = str(plosiveDict[w])
   outfileDict.write(w)
   outfileDict.write('\t')
   outfileDict.write(plosivecount)
   outfileDict.write('\n')
print ('\n', 'Total plosives: ', counter)
counter = str(counter)
outfileDict.write('\n')
outfileDict.write("Total plosives: ")
outfileDict.write(counter)
infile.close()
outfileDict.close()
```

9.8.2 Move files to different folder

```
#!/usr/bin/python
# -*- coding: utf-8 -*-
# Terminal : python3 moveFilestoFolder.py
# cwatte
# Script moves certain files to different folder

import re, codecs, os, sys
import shutil
path = '_ALLIntervalsOfPlosives_Sound_TXT/'
file_list = []
# collect all elements in folder in a list
for files in os.listdir(path):
   file_list.append(files)

txtCounter = 0
wavCounter = 0
for f in file_list:
   d1 = path + f
   if f.endswith("txt"):
      d2 = '_ALLIntervalsOfPlosives_TXT/' + f
      shutil.move(d1, d2)
      txtCounter = txtCounter + 1

   if f.endswith("wav"):
      d2 = '_ALLIntervalsOfPlosives_WAV/' + f
      shutil.move(d1, d2)
      wavCounter = wavCounter + 1
print ("txt: ", txtCounter, "\n", "wav: ", txtCounter)
```

9.8.3 Collect unique name of each plosive in corpus

```
#!/usr/bin/python
# -*- coding: utf-8 -*-
# Terminal : python3 getiD.py
# cwatte
# Script goes through all wav files in folder and collects plosive ID
# collect ID in table
# plosives ID = unique name of each plosive in corpus

import re, codecs, os, sys
path = '_ALLIntervalsOfPlosives_WAV/'
file_list = []

for files in os.listdir(path):
   file_list.append(files)
print (len(file_list))
```

```
file_list = sorted(file_list)
pCounter = 0
with codecs.open('id_of_Plosives.txt', 'w') as outfile:
   for f in file_list:
      pCounter += 1
      p= str(pCounter)
      idPlosive = f.rstrip(".wav")
      print(idPlosive)
      outfile.write(p)
      outfile.write('\t')
      outfile.write(idPlosive)
      outfile.write('\n')
```

9.8.4 Rename all plosives by giving each file unique number

```
#!/usr/bin/python
# -*- coding: utf-8 -*-
# Terminal : python3 getiD.py
# cwatte
# Script goes through files in directory and renames them
# by adding unique number for each plosive

import os
counter = 1
inputdir = "/Users/..."
outputdir = "/Users/..."
outfile = open("/Users/.../info_Renaming.txt","w")

for file in sorted(os.listdir(inputdir)):

   # make all counters 3 characters long
   strcounter = str(counter)
   if len(strcounter) == 1:
      strcounter = "00" + strcounter
      print (strcounter)
   if len(strcounter) == 2:
      strcounter = "0" + strcounter
      print (strcounter)

   newName = strcounter + "_" + file

   replacement = ".wav"
   idName = file.replace(replacement, "")
   nr_idName = newName.replace(replacement, "")

   outfile.write(strcounter)
   outfile.write("\t")
   outfile.write(idName)
   outfile.write("\t")
   outfile.write(nr_idName)
   outfile.write("\t")
   outfile.write(newName)
   outfile.write("\n")

   into = inputdir + "/" + file
   out = outputdir + "/" + newName
   counter += 1
outfile.close()
```

9.8.5 Save all selected files in Praat object window

```
# Script saves all selected objects from praat object window

# user defines target directory to save files
# form Select objects and define target directory
# sentence folder
# endform

# or define path here
folder$ = "/Users/..."
dir$ = folder$ + "/"
appendInfoLine: dir$

# store selected TGs
number_of_selected_TGs = numberOfSelected ("TextGrid")
#appendInfoLine: number_of_selected_TGs
for i to number_of_selected_TGs
   tg[i] = selected("TextGrid", i)
```

```
endfor

# store selected sounds
number_of_selected_sounds = numberOfSelected ("Sound")
for i to number_of_selected_sounds
   sound[i] = selected("Sound", i)
endfor

tgCounter = 1
appendInfoLine: "Saved TextGrids:"
# loop though stored TGs and sounds to save them in target directory
for i to number_of_selected_TGs
   selectObject: tg[i]
   name$ = selected$ ("TextGrid")
   appendInfoLine: tgCounter, tab$, name$
   Save as text file: dir$ + name$ + ".TextGrid"
   tgCounter = tgCounter + 1
endfor

appendInfoLine: ""
sCounter = 1
appendInfoLine: "Saved Sounds:"
for i to number_of_selected_sounds
   selectObject: sound[i]
   name$ = selected$ ("Sound")
   appendInfoLine: sCounter, tab$, name$
   Write to WAV file: dir$ + name$ + ".wav"
   sCounter = sCounter + 1
endfor
```

9.9 R Scripts

9.9.1 R script for statistical analysis and visualizations

```
# remove all objects from Global Environment
rm(list = ls())
library(dplyr)
library(ggplot2)
library("coin")

table_name <- "/Users/.../_DIAPIX_Korpus_csv.txt"
data <- read.csv(table_name, header= TRUE)

# Split Data in smaller datasets for analysis
data.Condensed <- data %>%
  dplyr::select(Speaker, Class, Gender, Aspiration, Plosive.Letter, Plosive, Preceding.Sound,
   Closure_ms, Release_ms, pVOT, Duration_TargetW.P_ms) %>%
  group_by(Class)
write.table(data.Condensed, file="/Users/.../dataCondensed_latex.txt", sep="&", quote=FALSE)

########################################
#### Absolute Numbers
########################################
### prepare tables for analysis
# display mean per group
tapply(data.Condensed$Release_ms, data.Condensed$Class, mean)

# create tables for PoA
data.p <- data.Condensed[data.Condensed$Plosive.Letter == 'p', ]
data.t <- data.Condensed[data.Condensed$Plosive.Letter == 't', ]
tapply(data.p$Release_ms, data.p$Class, mean)
tapply(data.t$Release_ms, data.t$Class, mean)

# tables for aspiration
data.p.A1 <- data.p[data.p$Aspiration == 1, ]
data.p.A0 <- data.p[data.p$Aspiration == 0, ]
data.t.A1 <- data.t[data.t$Aspiration == 1, ]
data.t.A0 <- data.t[data.t$Aspiration == 0, ]
tapply(data.p.A1$Release_ms, data.p.A1$Class, mean)
tapply(data.p.A0$Release_ms, data.p.A0$Class, mean)
tapply(data.t.A1$Release_ms, data.t.A1$Class, mean)
tapply(data.t.A0$Release_ms, data.t.A0$Class, mean)

# tables for gender
data.p.f <- data.p[data.p$Gender == "f", ]
data.p.m <- data.p[data.p$Gender == "m", ]
data.t.f <- data.t[data.t$Gender == "f", ]
```

```
data.t.m <- data.t[data.t$Gender == "m", ]
data.p.f.A <- data.p.f[data.p.f$Aspiration == "0",]
tapply(data.p.f$Release_ms, data.p.f$Class, mean)
tapply(data.p.m$Release_ms, data.p.m$Class, mean)
tapply(data.t.f$Release_ms, data.t.f$Class, mean)
tapply(data.t.m$Release_ms, data.t.m$Class, mean)

###
### Witihin Speaker variabiliy
###
data.mono <- data.Condensed[data.Condensed$Class == "mono", ]
data.multi <- data.Condensed[data.Condensed$Class == "multi", ]
data.o60 <- data.Condensed[data.Condensed$Class == "o60", ]
tapply(data.mono$Release_ms, data.mono$Speaker, mad)
tapply(data.multi$Release_ms, data.multi$Speaker, mad)
tapply(data.o60$Release_ms, data.o60$Speaker, mad)
tapply(data.mono$Release_ms, data.mono$Speaker, median)
tapply(data.multi$Release_ms, data.multi$Speaker, median)
tapply(data.o60$Release_ms, data.o60$Speaker, median)
tapply(data.Condensed$Release_ms, data.Condensed$Class, mad)
tapply(data.Condensed$Release_ms, data.Condensed$Class, median)
tapply(data.Condensed$Release_ms, data.Condensed$Class, mean)
tapply(data.Condensed$Release_ms, data.Condensed$Class, sd)

############
# density plot of all three groups together
############
median_VOT <- data.Condensed %>%
  group_by(Class) %>%
  summarize(median=median(Release_ms))
mean_VOT <- data.Condensed %>%
  group_by(Class) %>%
  summarize(mean=mean(Release_ms))

# density plot all groups
density.class <- ggplot(data.Condensed, aes(x=Release_ms, color=Class)) +
  xlim(0,130) +
  ylim(0,0.06) +
  ggtitle("Median per group") +
  xlab("VOT (ms)") +
  ylab("Density")+
  geom_density()+
  geom_vline(data=median_VOT, aes(xintercept=median, color=Class),
             linetype="dashed")
density.class

# density plot all groups
density.class <- ggplot(data.Condensed, aes(x=Release_ms, color=Class)) +
  xlim(0,130) +
  ylim(0,0.06) +
  ggtitle("Mean per group") +
  xlab("VOT (ms)") +
  ylab("Density")+
  geom_density()+
  geom_vline(data=mean_VOT, aes(xintercept=mean, color=Class),
             linetype="dashed")
density.class

# difference in aspiration
data.A1 <- data.Condensed[data.Condensed$Aspiration == "1", ]
data.A0 <- data.Condensed[data.Condensed$Aspiration == "0", ]
tapply(data.A1$Release_ms, data.A1$Class, mad)
tapply(data.A1$Release_ms, data.A1$Class, median)
tapply(data.A0$Release_ms, data.A0$Class, mad)
tapply(data.A0$Release_ms, data.A0$Class, median)

median_A1 <- data.A1 %>%
  group_by(Class) %>%
  summarize(median=median(Release_ms))
median_A0 <- data.A0 %>%
  group_by(Class) %>%
  summarize(median=median(Release_ms))

density.A1 <- ggplot(data.A1, aes(x=Release_ms, color=Class)) +
  xlim(0,130) +
  ylim(0,0.06) +
  ggtitle("Median per group [+Aspiration]") +
  xlab("VOT (ms)") +
  ylab("Density")+
  geom_density()+
  geom_vline(data=median_A1, aes(xintercept=median, color=Class),
```

```
             linetype="dashed")
density.A1

density.A0 <- ggplot(data.A0, aes(x=Release_ms, color=Class)) +
  xlim(0,130) +
  ylim(0,0.06) +
  ggtitle("Median per group [-Aspiration]") +
  xlab("VOT (ms)") +
  ylab("Density")+
  geom_density()+
  geom_vline(data=median_A0, aes(xintercept=median, color=Class),
             linetype="dashed")
density.A0

# separate group plots
density.mono <- ggplot(data.mono, aes(x=Release_ms, color=Speaker)) +
  xlim(0,120) +
  ylim(0,0.1) +
  ggtitle("mono") +
  xlab("VOT (ms)") +
  ylab("Density")+
  geom_density()
density.mono

density.multi <- ggplot(data.multi, aes(x=Release_ms, color=Speaker)) +
  xlim(0,120) +
  ylim(0,0.1) +
  ggtitle("multi") +
  xlab("VOT (ms)") +
  ylab("Density")+
  geom_density()
density.multi

density.o60 <- ggplot(data.o60, aes(x=Release_ms, color=Speaker)) +
  xlim(0,120) +
  ylim(0,0.1) +
  ggtitle("o60") +
  xlab("VOT (ms)") +
  ylab("Density")+
  geom_density()
density.o60

ggarrange(density.mono, density.multi, density.o60 + rremove("x.text"),
          labels = c("", "", ""),
          ncol = 2, nrow = 2)

boxplot(data.mono$Release_ms ~ data.mono$Speaker,
                      # names =c("mono", "multi"),
                      las=1,
                      #horizontal = TRUE,
                      main = "Mono",
                      ylim = c(0, 120),
                      ylab = "Time (ms)",
                      xlab = "")
boxplot(data.multi$Release_ms ~ data.multi$Speaker,
        # names =c("mono", "multi"),
        las=1,
        #horizontal = TRUE,
        #ylim = c(0,120)),
        main = "Multi",
        ylim = c(0, 120),
        ylab = "",
        xlab = "")
boxplot(data.o60$Release_ms ~ data.o60$Speaker,
        # names =c("mono", "multi"),
        las=1,
        #horizontal = TRUE,
        main = "o60",
        ylim = c(0, 120),
        ylab = "",
        xlab = "")

### Visualize means
# Group
boxplot(data.Condensed$Release_ms ~ data.Condensed$Class,
        xlab = "Class",
        ylab = "VOT (ms)",
        main = "Mean VOT per group")
tapply(data.Condensed$Release_ms, data.Condensed$Class, mean)
tapply(data.Condensed$Release_ms, data.Condensed$Class, sd)
tapply(data.Condensed$Release_ms, data.Condensed$Class, mad)
```

```
boxplot_class <- ggplot(data.Condensed, aes(Class, Release_ms)) +
  geom_boxplot() +
  geom_jitter(color="black", size=0.4, alpha=0.9) +
  #geom_hline(yintercept =14.83, linetype="solid", color = "red", size=0.5) +
  #geom_segment(x= )
  #geom_line(x=data.Condensed$Class=="mono", y = 14) +
  labs(x="", y="Time (ms)") +
  labs(title = paste("Mean VOT /p/ and /t/ per Class"))
plot(boxplot_class)

# PoA
boxplot(data.p$Release_ms ~ data.p$Class,
        xlab = "",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "/p/")
boxplot(data.t$Release_ms ~ data.t$Class,
        xlab = "",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "/t/")

# Aspiration
boxplot(data.p.A1$Release_ms ~ data.p.A1$Class,
        xlab = "",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "[ph]")
boxplot(data.p.A0$Release_ms ~ data.p.A0$Class,
        xlab = "",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "[p]")
boxplot(data.t.A1$Release_ms ~ data.t.A1$Class,
        xlab = "",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "[th]")
boxplot(data.t.A0$Release_ms ~ data.t.A0$Class,
        xlab = "",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "[t]")

# Gender
boxplot(data.p.f$Release_ms ~ data.p.f$Class,
        xlab = "",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "f /p/")
boxplot(data.p.m$Release_ms ~ data.p.m$Class,
        xlab = "",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "m /p/")
boxplot(data.t.f$Release_ms ~ data.t.f$Class,
        xlab = "",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "f /t/")
boxplot(data.t.m$Release_ms ~ data.t.m$Class,
        xlab = "",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "m /t/")

###
### NORMALITY TEST: SHAPIRO TEST + QQplots + Density plots
### if p-value < 0.05 => sample not normally distributed
# Group
aggregate(Release_ms ~ Class, data = data.Condensed, function(x) shapiro.test(x)$p.value)
nrow(data.Condensed[data.Condensed$Class == "mono",]) # 163
nrow(data.Condensed[data.Condensed$Class == "multi",]) # 120
nrow(data.Condensed[data.Condensed$Class == "o60",]) # 243
data.mono <- data.Condensed[data.Condensed$Class == 'mono', ]
qqPlot(data.mono$Release_ms,
       main="VOT group mono",
       ylab="VOT (ms)")
ggdensity(data.mono$Release_ms,
          main = "VOT group mono",
```

```
          ylim = c(0,0.06),
          xlim= c(0,120),
          xlab = "VOT (ms)")

data.multi <- data.Condensed[data.Condensed$Class == 'multi', ]
qqPlot(data.multi$Release_ms,
       main="VOT group multi",
       ylab="VOT (ms)")
ggdensity(data.multi$Release_ms,
          main = "VOT group multi",
          ylim = c(0,0.06),
          xlim= c(0,120),
          xlab = "VOT (ms)")

data.o60 <- data.Condensed[data.Condensed$Class == 'o60', ]
qqPlot(data.o60$Release_ms,
       main="VOT group o60",
       ylab="VOT (ms)")
ggdensity(data.o60$Release_ms,
          main = "VOT group o60",
          ylim = c(0,0.06),
          xlim= c(0,120),
          xlab = "VOT (ms)")

#%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ggplot + density required pacakges : car, ggpur, ggplot

# PoA
# ------------------- plosive /p/
aggregate(Release_ms ~ Class, data = data.p, function(x) shapiro.test(x)$p.value)
qqPlot(data.p$Release_ms) #
nrow(data.p[data.p$Class == "mono",]) # 48
nrow(data.p[data.p$Class == "multi",]) # 38
nrow(data.p[data.p$Class == "o60",]) # 33
data.mono.p <- data.mono[data.mono$Plosive.Letter == 'p', ]
qqPlot(data.mono.p$Release_ms,
       main="VOT plosive /p/ mono",
       ylab="")
ggdensity(data.mono.p$Release_ms,
          main = "VOT plosive /p/ mono",
          xlab = "VOT")
data.multi.p <- data.multi[data.multi$Plosive.Letter == 'p', ]
qqPlot(data.multi.p$Release_ms,
       main="VOT plosive /p/ multi",
       ylab="")
ggdensity(data.multi.p$Release_ms,
          main = "VOT plosive /p/ multi",
          xlab = "VOT")
data.o60.p <- data.o60[data.o60$Plosive.Letter == 'p', ]
qqPlot(data.o60.p$Release_ms,
       main="VOT plosive /p/ o60",
       ylab="")
ggdensity(data.o60.p$Release_ms,
          main = "VOT plosive /p/ o60",
          xlab = "VOT")

# ------------------- plosive /t/
aggregate(Release_ms ~ Class, data = data.t, function(x) shapiro.test(x)$p.value)
qqPlot(data.t$Release_ms) # not normally distributed
nrow(data.t[data.t$Class == "mono",]) # 115
nrow(data.t[data.t$Class == "multi",]) # 82
nrow(data.t[data.t$Class == "o60",]) # 210
data.mono.t <- data.mono[data.mono$Plosive.Letter == 't', ]
qqPlot(data.mono.t$Release_ms,
       main="VOT plosive /t/ mono",
       ylab="")
ggdensity(data.mono.t$Release_ms,
          main = "VOT plosive /t/ mono",
          xlab = "VOT")
data.multi.t <- data.multi[data.multi$Plosive.Letter == 't', ]
qqPlot(data.multi.t$Release_ms,
       main="VOT plosive /t/ multi",
       ylab="")
ggdensity(data.multi.t$Release_ms,
          main = "VOT plosive /t/ multi",
          xlab = "VOT")
data.o60.t <- data.o60[data.o60$Plosive.Letter == 't', ]
qqPlot(data.o60.t$Release_ms,
       main="VOT plosive /t/ o60",
       ylab="")
ggdensity(data.o60.t$Release_ms,
          main = "VOT plosive /t/ o60",
```

```
          xlab = "VOT")

################
# aspiration
# ---------------- aspiration [ph]
aggregate(Release_ms ~ Class, data = data.p.A1, function(x) shapiro.test(x)$p.value)
qqPlot(data.p.A1$Release_ms)
nrow(data.p.A1[data.p.A1$Class == "mono",]) # 41
nrow(data.p.A1[data.p.A1$Class == "multi",]) # 25
nrow(data.p.A1[data.p.A1$Class == "o60",]) # 13
data.mono.p.A1 <- data.mono.p[data.mono.p$Aspiration == '1', ]
qqPlot(data.mono.p.A1$Release_ms,
       main="VOT aspiration [ph] mono",
       ylab="")
ggdensity(data.mono.p.A1$Release_ms,
          main = "VOT aspiration [ph] mono",
          xlab = "VOT")
data.multi.p.A1 <- data.multi.p[data.multi.p$Aspiration == '1', ]
qqPlot(data.multi.p.A1$Release_ms,
       main="VOT aspiration [ph] multi",
       ylab="")
ggdensity(data.multi.p.A1$Release_ms,
          main = "VOT aspiration [ph] multi",
          xlab = "VOT")
data.o60.p.A1 <- data.o60.p[data.o60.p$Aspiration == '1', ]
qqPlot(data.o60.p.A1$Release_ms,
       main="VOT aspiration [ph] o60",
       ylab="")
ggdensity(data.o60.p.A1$Release_ms,
          main = "VOT aspiration [ph] o60",
          xlab = "VOT")

# ---------------- aspiration [p]
aggregate(Release_ms ~ Class, data = data.p.A0, function(x) shapiro.test(x)$p.value)
qqPlot(data.p.A0$Release_ms) # normal
nrow(data.p.A0[data.p.A0$Class == "mono",]) # 7
nrow(data.p.A0[data.p.A0$Class == "multi",]) # 13
nrow(data.p.A0[data.p.A0$Class == "o60",]) # 20
data.mono.p.A0 <- data.mono.p[data.mono.p$Aspiration == '0', ]
qqPlot(data.mono.p.A0$Release_ms,
       main="VOT aspiration [p] mono",
       ylab="")
ggdensity(data.mono.p.A0$Release_ms,
          main = "VOT aspiration [p] mono",
          xlab = "VOT")
data.multi.p.A0 <- data.multi.p[data.multi.p$Aspiration == '0', ]
qqPlot(data.multi.p.A0$Release_ms,
       main="VOT aspiration [p] multi",
       ylab="")
ggdensity(data.multi.p.A0$Release_ms,
          main = "VOT aspiration [p] multi",
          xlab = "VOT")
data.o60.p.A0 <- data.o60.p[data.o60.p$Aspiration == '0', ]
qqPlot(data.o60.p.A0$Release_ms,
       main="VOT aspiration [p] o60",
       ylab="")
ggdensity(data.o60.p.A0$Release_ms,
          main = "VOT aspiration [p] o60",
          xlab = "VOT")

# ---------------- aspiration [th]
aggregate(Release_ms ~ Class, data = data.t.A1, function(x) shapiro.test(x)$p.value)
qqPlot(data.t.A1$Release_ms) # normal
nrow(data.t.A1[data.t.A1$Class == "mono",]) # 12
nrow(data.t.A1[data.t.A1$Class == "multi",]) # 9
nrow(data.t.A1[data.t.A1$Class == "o60",]) # 10
data.mono.t.A1 <- data.mono.t[data.mono.t$Aspiration == '1', ]
qqPlot(data.mono.t.A1$Release_ms,
       main="VOT aspiration [th] mono",
       ylab="")
ggdensity(data.mono.t.A1$Release_ms,
          main = "VOT aspiration [th] mono",
          xlab = "VOT")
data.multi.t.A1 <- data.multi.t[data.multi.t$Aspiration == '1', ]
qqPlot(data.multi.t.A1$Release_ms,
       main="VOT aspiration [th] multi",
       ylab="")
ggdensity(data.multi.t.A1$Release_ms,
          main = "VOT aspiration [th] multi",
          xlab = "VOT")
data.o60.t.A1 <- data.o60.t[data.o60.t$Aspiration == '1', ]
```

```
qqPlot(data.o60.t.A1$Release_ms,
       main="VOT aspiration [th] o60",
       ylab="")
ggdensity(data.o60.t.A1$Release_ms,
          main = "VOT aspiration [th] o60",
          xlab = "VOT")

# ---------------- aspiration [t]
aggregate(Release_ms ~ Class, data = data.t.A0, function(x) shapiro.test(x)$p.value)
qqPlot(data.t.A0$Release_ms) # not normally distributed
nrow(data.t.A0[data.t.A0$Class == "mono",]) # 103
nrow(data.t.A0[data.t.A0$Class == "multi",]) # 73
nrow(data.t.A0[data.t.A0$Class == "o60",]) # 200
data.mono.t.A0 <- data.mono.t[data.mono.t$Aspiration == '0', ]
qqPlot(data.mono.t.A0$Release_ms,
       main="VOT aspiration [t] mono",
       ylab="")
ggdensity(data.mono.t.A0$Release_ms,
          main = "VOT aspiration [t] mono",
          xlab = "VOT")
data.multi.t.A0 <- data.multi.t[data.multi.t$Aspiration == '0', ]
qqPlot(data.multi.t.A0$Release_ms,
       main="VOT aspiration [t] multi",
       ylab="")
ggdensity(data.multi.t.A0$Release_ms,
          main = "VOT aspiration [t] multi",
          xlab = "VOT")
data.o60.t.A0 <- data.o60.t[data.o60.t$Aspiration == '0', ]
qqPlot(data.o60.t.A0$Release_ms,
       main="VOT aspiration [t] o60",
       ylab="")
ggdensity(data.o60.t.A0$Release_ms,
          main = "VOT aspiration [t] o60",
          xlab = "VOT")

# Gender
# ---------------- Gender /p/ f
aggregate(Release_ms ~ Class, data = data.p.f, function(x) shapiro.test(x)$p.value)
qqPlot(data.p.f$Release_ms) # not normally distributed
nrow(data.p.f[data.p.f$Class == "mono",]) # 33
nrow(data.p.f[data.p.f$Class == "multi",]) # 19
nrow(data.p.f[data.p.f$Class == "o60",]) # 24
data.mono.p.f <- data.mono.p[data.mono.p$Gender == 'f', ]
qqPlot(data.mono.p.f$Release_ms,
       main="VOT gender /p/ f mono",
       ylab="")
ggdensity(data.mono.p.f$Release_ms,
          main = "VOT gender /p/ f mono",
          xlab = "VOT")
data.multi.p.f <- data.multi.p[data.multi.p$Gender == 'f', ]
qqPlot(data.multi.p.f$Release_ms,
       main="VOT gender /p/ f multi",
       ylab="")
ggdensity(data.multi.p.f$Release_ms,
          main = "VOT gender /p/ f multi",
          xlab = "VOT")
data.o60.p.f <- data.o60.p[data.o60.p$Gender == 'f', ]
qqPlot(data.o60.p.f$Release_ms,
       main="VOT gender /p/ f o60",
       ylab="")
ggdensity(data.o60.p.f$Release_ms,
          main = "VOT gender /p/ f o60",
          xlab = "VOT")

# ---------------- Gender /p/ m
aggregate(Release_ms ~ Class, data = data.p.m, function(x) shapiro.test(x)$p.value)
qqPlot(data.p.m$Release_ms) # normal
nrow(data.p.m[data.p.m$Class == "mono",]) # 15
nrow(data.p.m[data.p.m$Class == "multi",]) # 19
nrow(data.p.m[data.p.m$Class == "o60",]) # 9
data.mono.p.m <- data.mono.p[data.mono.p$Gender == 'm', ]
qqPlot(data.mono.p.m$Release_ms,
       main="VOT gender /p/ m mono",
       ylab="")
ggdensity(data.mono.p.m$Release_ms,
          main = "VOT gender /p/ m mono",
          xlab = "VOT")
data.multi.p.m <- data.multi.p[data.multi.p$Gender == 'm', ]
qqPlot(data.multi.p.m$Release_ms,
       main="VOT gender /p/ m multi",
       ylab="")
```

```
ggdensity(data.multi.p.m$Release_ms,
          main = "VOT gender /p/ m multi",
          xlab = "VOT")
data.o60.p.m <- data.o60.p[data.o60.p$Gender == 'm', ]
qqPlot(data.o60.p.m$Release_ms,
       main="VOT gender /p/ m o60",
       ylab="")
ggdensity(data.o60.p.m$Release_ms,
          main = "VOT gender /p/ m o60",
          xlab = "VOT")

# ----------------- Gender /t/ f
aggregate(Release_ms ~ Class, data = data.t.f, function(x) shapiro.test(x)$p.value)
qqPlot(data.t.f$Release_ms)
nrow(data.t.f[data.t.f$Class == "mono",]) # 86
nrow(data.t.f[data.t.f$Class == "multi",]) # 24
nrow(data.t.f[data.t.f$Class == "o60",]) # 120
data.mono.t.f <- data.mono.t[data.mono.t$Gender == 'f', ]
qqPlot(data.mono.t.f$Release_ms,
       main="VOT gender /t/ f mono",
       ylab="")
ggdensity(data.mono.t.f$Release_ms,
          main = "VOT gender /t/ f mono",
          xlab = "VOT")
data.multi.t.f <- data.multi.t[data.multi.t$Gender == 'f', ]
qqPlot(data.multi.t.f$Release_ms,
       main="VOT gender /t/ f multi",
       ylab="")
ggdensity(data.multi.t.f$Release_ms,
          main = "VOT gender /t/ f multi",
          xlab = "VOT")
data.o60.t.f <- data.o60.t[data.o60.t$Gender == 'f', ]
qqPlot(data.o60.t.f$Release_ms,
       main="VOT gender /t/ f o60",
       ylab="")
ggdensity(data.o60.t.f$Release_ms,
          main = "VOT gender /t/ f o60",
          xlab = "VOT")

# ----------------- Gender /t/ m
aggregate(Release_ms ~ Class, data = data.t.m, function(x) shapiro.test(x)$p.value)
qqPlot(data.t.m$Release_ms)  # not normally distributed
nrow(data.t.m[data.t.m$Class == "mono",]) # 29
nrow(data.t.m[data.t.m$Class == "multi",]) # 58
nrow(data.t.m[data.t.m$Class == "o60",]) # 90
data.mono.t.m <- data.mono.t[data.mono.t$Gender == 'm', ]
qqPlot(data.mono.t.m$Release_ms,
       main="VOT gender /t/ m mono",
       ylab="")
ggdensity(data.mono.t.m$Release_ms,
          main = "VOT gender /t/ m mono",
          xlab = "VOT")
data.multi.t.m <- data.multi.t[data.multi.t$Gender == 'm', ]
qqPlot(data.multi.t.m$Release_ms,
       main="VOT gender /t/ m multi",
       ylab="")
ggdensity(data.multi.t.m$Release_ms,
          main = "VOT gender /t/ m multi",
          xlab = "VOT")
data.o60.t.m <- data.o60.t[data.o60.t$Gender == 'm', ]
qqPlot(data.o60.t.m$Release_ms,
       main="VOT gender /t/ m o60",
       ylab="")
ggdensity(data.o60.t.m$Release_ms,
          main = "VOT gender /t/ m o60",
          xlab = "VOT")
###
### IQR, mad and SD
# Class
tapply(data.Condensed$Release_ms, data.Condensed$Class, sd)
tapply(data.Condensed$Release_ms, data.Condensed$Class, IQR)
tapply(data.Condensed$Release_ms, data.Condensed$Class, mad)

# PoA
tapply(data.p$Release_ms, data.p$Class, sd)
tapply(data.p$Release_ms, data.p$Class, IQR)
tapply(data.p$Release_ms, data.p$Class, mad)
tapply(data.t$Release_ms, data.t$Class, sd)
tapply(data.t$Release_ms, data.t$Class, IQR)
tapply(data.t$Release_ms, data.t$Class, mad)
```

```
# Aspiration
tapply(data.p.A1$Release_ms, data.p.A1$Class, sd)
tapply(data.p.A1$Release_ms, data.p.A1$Class, IQR)
tapply(data.p.A1$Release_ms, data.p.A1$Class, mad)
tapply(data.p.A0$Release_ms, data.p.A0$Class, sd)
tapply(data.p.A0$Release_ms, data.p.A0$Class, IQR)
tapply(data.p.A0$Release_ms, data.p.A0$Class, mad)
tapply(data.t.A1$Release_ms, data.t.A1$Class, sd)
tapply(data.t.A1$Release_ms, data.t.A1$Class, IQR)
tapply(data.t.A1$Release_ms, data.t.A1$Class, mad)
tapply(data.t.A0$Release_ms, data.t.A0$Class, sd)
tapply(data.t.A0$Release_ms, data.t.A0$Class, IQR)
tapply(data.t.A0$Release_ms, data.t.A0$Class, mad)

# Gender
tapply(data.p.f$Release_ms, data.p.f$Class, sd)
tapply(data.p.f$Release_ms, data.p.f$Class, IQR)
tapply(data.p.f$Release_ms, data.p.f$Class, mad)
tapply(data.p.m$Release_ms, data.p.m$Class, sd)
tapply(data.p.m$Release_ms, data.p.m$Class, IQR)
tapply(data.p.m$Release_ms, data.p.m$Class, mad)
tapply(data.t.f$Release_ms, data.t.f$Class, sd)
tapply(data.t.f$Release_ms, data.t.f$Class, IQR)
tapply(data.t.f$Release_ms, data.t.f$Class, mad)
tapply(data.t.m$Release_ms, data.t.m$Class, sd)
tapply(data.t.m$Release_ms, data.t.m$Class, IQR)
tapply(data.t.m$Release_ms, data.t.m$Class, mad)

###
### TEST HOMOGENEITY OF VARIANCE
###
# class
leveneTest(Release_ms ~ Class, data = data.Condensed)
bartlett.test(Release_ms ~ Class, data = data.Condensed) # no : p-value = 0.0005872
fligner.test(Release_ms ~ Class, data = data.Condensed) # no: p-value = 1.214e-08

# PoA
bartlett.test(Release_ms ~ Class, data = data.p) # yes : p-value = 0.2088
fligner.test(Release_ms ~ Class, data = data.p) # yes : p-value = 0.5577
bartlett.test(Release_ms ~ Class, data = data.t) # no : p-value = 7.343e-08
fligner.test(Release_ms ~ Class, data = data.t) # no : p-value = 7.024e-12

# aspiration
bartlett.test(Release_ms ~ Class, data = data.p.A1) # yes : p-value = 0.5128
fligner.test(Release_ms ~ Class, data = data.p.A1) # yes : p-value = 0.4645
bartlett.test(Release_ms ~ Class, data = data.p.A0) # no : p-value = 0.03872
fligner.test(Release_ms ~ Class, data = data.p.A0) # yes : p-value = 0.1729
bartlett.test(Release_ms ~ Class, data = data.t.A1) # no : p-value = 0.04378
fligner.test(Release_ms ~ Class, data = data.t.A1) # yes : p-value = 0.5196
bartlett.test(Release_ms ~ Class, data = data.t.A0) # no : p-value < 2.2e-16
fligner.test(Release_ms ~ Class, data = data.t.A0) # no : p-value = 5.799e-16

# gender
bartlett.test(Release_ms ~ Class, data = data.p.f) # yes : p-value = 0.1397
fligner.test(Release_ms ~ Class, data = data.p.f) # yes : p-value = 0.244
bartlett.test(Release_ms ~ Class, data = data.p.m) # no : p-value = 0.02068
fligner.test(Release_ms ~ Class, data = data.p.m) # no : p-value = 0.02909
bartlett.test(Release_ms ~ Class, data = data.t.f) # no : p-value = 1.16e-06
fligner.test(Release_ms ~ Class, data = data.t.f) # no : p-value = 2.369e-07
bartlett.test(Release_ms ~ Class, data = data.t.m) # no : p-value = 0.0245
fligner.test(Release_ms ~ Class, data = data.t.m) # no : p-value = 4.945e-05

# ###
# ### ANOVA test if difference of groups are statistically different
# ###
### CLASS
# one_way()
npar.class <- nparcomp(Release_ms ~ Class, data = data.Condensed, type = "Tukey")
npar.class$Analysis
summary(npar.class)
plot(npar.class)

### PoA
### /p/
kruskal.test(Release_ms ~ Class, data = data.p)
npar.p <- nparcomp(Release_ms ~ Class, data = data.p, type = "Tukey")
npar.p$Analysis
plot(npar.p)

### /t/
# one_way()
```

```
npar.t <- nparcomp(Release_ms ~ Class, data = data.t, type = "Tukey")
npar.t$Analysis
plot(npar.t)

### ASPIRATION
### /p/ A1
kruskal.test(Release_ms ~ Class, data = data.p.A1)
npar.p.A1 <- nparcomp(Release_ms ~ Class, data = data.p.A1, type = "Tukey")
npar.p.A1$Analysis

### /p/ A0
# one_way()
npar.p.A0 <- nparcomp(Release_ms ~ Class, data = data.p.A0, type = "Tukey")
npar.p.A0$Analysis

### /t/ A1
oneway.test(Release_ms ~ Class, data = data.t.A1)
npar.t.A1 <- nparcomp(Release_ms ~ Class, data = data.t.A1, type = "Tukey")
npar.t.A1$Analysis

### /t/ A0
# one_way()
npar.t.A0 <- nparcomp(Release_ms ~ Class, data = data.t.A0, type = "Tukey")
npar.t.A0$Analysis

### GENDER
### /p/ f
kruskal.test(Release_ms ~ Class, data = data.p.f)
npar.p.f <- nparcomp(Release_ms ~ Class, data = data.p.f, type = "Tukey")
npar.p.f$Analysis

### /p/ m
oneway.test(Release_ms ~ Class, data = data.p.m)
npar.p.m <- nparcomp(Release_ms ~ Class, data = data.p.m, type = "Tukey")
npar.p.m$Analysis

### /t/ f
# one_way()
npar.t.f <- nparcomp(Release_ms ~ Class, data = data.t.f, type = "Tukey")
npar.t.f$Analysis

### /t/ m
# one_way()
npar.t.m <- nparcomp(Release_ms ~ Class, data = data.t.m, type = "Tukey")
npar.t.m$Analysis

###########
# Prepare data for oneway_test()
# convert class to factor
# mono = 1; multi = 2; o60=3
##############
data.Condensed.1way <- data.Condensed %>%
  dplyr::select(Speaker, Class, Gender, Aspiration, Plosive.Letter, Release_ms) %>%
  group_by(Class)

for (i in 1:526){
  if (data.Condensed[i, "Class"] == "mono"){
    data.Condensed.1way[i, "Class"] <- "1"
  }
  else if (data.Condensed[i, "Class"] == "multi"){
    data.Condensed.1way[i, "Class"] <- "2"
  }
  else if (data.Condensed[i, "Class"] == "o60"){
    data.Condensed.1way[i, "Class"] <- "3"
  }
}

data.Condensed.1way$Class <- as.factor(data.Condensed.1way$Class)

# CLASS
oneway_test(Release_ms ~ Class, data = data.Condensed.1way)

# PoA /t/
data.t.1way <- data.Condensed.1way[data.Condensed.1way$Plosive.Letter == 't', ]
oneway_test(Release_ms ~ Class, data = data.t.1way)

# ASPIRATION [t]
data.t.A0.1way <- data.t.1way[data.t.1way$Aspiration == '0', ]
oneway_test(Release_ms ~ Class, data = data.t.A0.1way)

# GENDER /t/ F + m
```

```
data.t.f.1way <- data.t.1way[data.t.1way$Gender == 'f', ]
data.t.m.1way <- data.t.1way[data.t.1way$Gender == 'm', ]
oneway_test(Release_ms ~ Class, data = data.t.f.1way)
oneway_test(Release_ms ~ Class, data = data.t.m.1way)

########################
# non-parametric Wilcoxon or Matt- Whitney tests.
# test for differences between p and ph / t and th in each class
########################
data.p.mono <- data.p[data.p$Class == "mono", ]
data.p.multi <- data.p[data.p$Class == "multi", ]
data.p.o60 <- data.p[data.p$Class == "o60", ]
data.t.mono <- data.t[data.t$Class == "mono", ]
data.t.multi <- data.t[data.t$Class == "multi", ]
data.t.o60 <- data.t[data.t$Class == "o60", ]

boxplot(data.p.mono$Release_ms ~ data.p.mono$Aspiration,
        xlab = "Aspiration",
        ylab = "VOT (ms)",
        ylim = c(0,100),
        main = "Mono: [p] vs [ph]")
boxplot(data.p.multi$Release_ms ~ data.p.multi$Aspiration,
        xlab = "Aspiration",
        ylab = "VOT (ms)",
        ylim = c(0,100),
        main = "Multi: [p] vs [ph]")
boxplot(data.p.o60$Release_ms ~ data.p.o60$Aspiration,
        xlab = "Aspiration",
        ylab = "VOT (ms)",
        ylim = c(0,100),
        main = "o60: [p] vs [ph]")
boxplot(data.t.mono$Release_ms ~ data.t.mono$Aspiration,
        xlab = "Aspiration",
        ylab = "VOT (ms)",
        ylim = c(0,100),
        main = "Mono: [t] vs [th]")
boxplot(data.t.multi$Release_ms ~ data.t.multi$Aspiration,
        xlab = "Aspiration",
        ylab = "VOT (ms)",
        ylim = c(0,100),
        main = "Multi: [t] vs [th]")
boxplot(data.t.o60$Release_ms ~ data.t.o60$Aspiration,
        xlab = "Aspiration",
        ylab = "VOT (ms)",
        ylim = c(0,100),
        main = "o60: [t] vs [th]")
fligner.test(Release_ms ~ Aspiration, data = data.p.mono) # yes : p-value = 0.5348
fligner.test(Release_ms ~ Aspiration, data = data.p.multi) # yes : p-value = 0.8788
fligner.test(Release_ms ~ Aspiration, data = data.p.o60) # no : p-value = 0.01764
bartlett.test(Release_ms ~ Aspiration, data = data.p.o60) # no : p-value = 0.0003718
fligner.test(Release_ms ~ Aspiration, data = data.t.mono) # yes : p-value = 0.9373
fligner.test(Release_ms ~ Aspiration, data = data.t.multi) # yes : p-value = 0.352
fligner.test(Release_ms ~ Aspiration, data = data.t.o60) # no : p-value = 0.6382
qqline(data.p.mono$Aspiration)
plot(density(data.p.mono$Aspiration))
data.p.mono.A1 <- data.p.mono[data.p.mono$Aspiration == "1", ]
data.p.mono.A0 <- data.p.mono[data.p.mono$Aspiration == "0", ]
wilcox.test(data.p.mono.A1$Release_ms, data.p.mono.A0$Release_ms, correct = FALSE, conf.int = TRUE)
data.p.multi.A1 <- data.p.multi[data.p.multi$Aspiration == "1", ]
data.p.multi.A0 <- data.p.multi[data.p.multi$Aspiration == "0", ]
wilcox.test(data.p.multi.A1$Release_ms, data.p.multi.A0$Release_ms, correct = FALSE, conf.int = TRUE)
data.p.o60.A1 <- data.p.o60[data.p.o60$Aspiration == "1", ]
data.p.o60.A0 <- data.p.o60[data.p.o60$Aspiration == "0", ]
wilcox.test(data.p.o60.A1$Release_ms, data.p.o60.A0$Release_ms, correct = FALSE, conf.int = TRUE)
data.t.mono.A1 <- data.t.mono[data.t.mono$Aspiration == "1", ]
data.t.mono.A0 <- data.t.mono[data.t.mono$Aspiration == "0", ]
wilcox.test(data.t.mono.A1$Release_ms, data.t.mono.A0$Release_ms, correct = FALSE, conf.int = TRUE)
data.t.multi.A1 <- data.t.multi[data.t.multi$Aspiration == "1", ]
data.t.multi.A0 <- data.t.multi[data.t.multi$Aspiration == "0", ]
wilcox.test(data.t.multi.A1$Release_ms, data.t.multi.A0$Release_ms, correct = FALSE, conf.int = TRUE)
data.t.o60.A1 <- data.t.o60[data.t.o60$Aspiration == "1", ]
data.t.o60.A0 <- data.t.o60[data.t.o60$Aspiration == "0", ]
wilcox.test(data.t.o60.A1$Release_ms, data.t.o60.A0$Release_ms, correct = FALSE, conf.int = TRUE)

########################
# non-parametric Wilcoxon or Matt- Whitney tests.
# test for differences between GENDER for /p/ and /t/ in each class
########################
data.p.mono.f <- data.p.mono[data.p.mono$Gender == "f", ]
data.p.mono.m <- data.p.mono[data.p.mono$Gender == "m", ]
wilcox.test(data.p.mono.f$Release_ms, data.p.mono.m$Release_ms, correct = FALSE, conf.int = TRUE)
```

```
data.t.mono.f <- data.t.mono[data.t.mono$Gender == "f", ]
data.t.mono.m <- data.t.mono[data.t.mono$Gender == "m", ]
wilcox.test(data.t.mono.f$Release_ms, data.t.mono.m$Release_ms, correct = FALSE, conf.int = TRUE)
data.p.multi.f <- data.p.multi[data.p.multi$Gender == "f", ]
data.p.multi.m <- data.p.multi[data.p.multi$Gender == "m", ]
wilcox.test(data.p.multi.f$Release_ms, data.p.multi.m$Release_ms, correct = FALSE, conf.int = TRUE)
data.t.multi.f <- data.t.multi[data.t.multi$Gender == "f", ]
data.t.multi.m <- data.t.multi[data.t.multi$Gender == "m", ]
wilcox.test(data.t.multi.f$Release_ms, data.t.multi.m$Release_ms, correct = FALSE, conf.int = TRUE)
data.p.o60.f <- data.p.o60[data.p.o60$Gender == "f", ]
data.p.o60.m <- data.p.o60[data.p.o60$Gender == "m", ]
wilcox.test(data.p.o60.f$Release_ms, data.p.o60.m$Release_ms, correct = FALSE, conf.int = TRUE)
data.t.o60.f <- data.t.o60[data.t.o60$Gender == "f", ]
data.t.o60.m <- data.t.o60[data.t.o60$Gender == "m", ]
wilcox.test(data.t.o60.f$Release_ms, data.t.o60.m$Release_ms, correct = FALSE, conf.int = TRUE)

boxplot(data.p.mono$Release_ms ~ data.p.mono$Gender,
        xlab = "Gender",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "Mono: /p/ f vs m")
boxplot(data.p.multi$Release_ms ~ data.p.multi$Gender,
        xlab = "Gender",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "Mulit: /p/ f vs m")
boxplot(data.p.o60$Release_ms ~ data.p.o60$Gender,
        xlab = "Gender",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "o60: /p/ f vs m")
boxplot(data.t.mono$Release_ms ~ data.t.mono$Gender,
        xlab = "Gender",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "Mono: /t/ f vs m")
boxplot(data.t.multi$Release_ms ~ data.t.multi$Gender,
        xlab = "Gender",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "Multi: /t/ f vs m")
boxplot(data.t.o60$Release_ms ~ data.t.o60$Gender,
        xlab = "Gender",
        ylab = "VOT (ms)",
        ylim = c(0,120),
        main = "o60: /t/ f vs m")
```

9.9.2 R script pVOT, further analysis, and visualizations

```
rm(list = ls())
library(dplyr)
library(ggplot2)

table_name <- "/Users/.../_DIAPIX_Korpus_csv.txt"
data <- read.csv(table_name, header= TRUE)

table_name_SYL <- "/Users/.../DIAPIX_Plosive_Korpus_SYL_csv.txt"
dataSYL <- read.csv(table_name_SYL, header= TRUE)
data <- dataSYL

### Split Data in smaller datasets for analysis
data.Condensed <- data %>%
  dplyr::select(Speaker, Class, Gender, Syllables, Aspiration, Plosive.Letter, Plosive, Preceding.Sound,
  Closure_ms, Release_ms, pVOT, Duration_TargetW.P_ms) %>%
  group_by(Class)

###########
# Prepare data for oneway_test()
# convert class to factor
# mono = 1; multi = 2; o60=3
##############
data.Condensed.1way <- data.Condensed %>%
  dplyr::select(Speaker, Class, Gender, Syllables, Aspiration, Plosive.Letter, Plosive, Release_ms, pVOT) %>%
    group_by(Class)
for (i in 1:526){
  if (data.Condensed[i, "Class"] == "mono"){
    data.Condensed.1way[i, "Class"] <- "1"
  }
  else if (data.Condensed[i, "Class"] == "multi"){
    data.Condensed.1way[i, "Class"] <- "2"
```

```
  }
  else if (data.Condensed[i, "Class"] == "o60"){
    data.Condensed.1way[i, "Class"] <- "3"
  }
}
data.Condensed.1way$Class <- as.factor(data.Condensed.1way$Class)

# ########################################
# Norm by syl
# ########################################
syl1 <-dataSYL[dataSYL$Syllables == '1', ]
tapply(syl1$pVOT, syl1$Class, mean)
tapply(syl1$pVOT, syl1$Class, sd)
tapply(syl1$pVOT, syl1$Class, median)
tapply(syl1$pVOT, syl1$Class, mad)
tapply(syl1$Release_ms, syl1$Class, mean)
tapply(syl1$Release_ms, syl1$Class, sd)
tapply(syl1$Release_ms, syl1$Class, median)
tapply(syl1$Release_ms, syl1$Class, mad)

boxplot(syl1$pVOT ~ syl1$Class,
        xlab = "Class",
        ylab = "Time (ms)",
        main = "VOT for one syllable words per Class")
aggregate(pVOT ~ Class, data = syl1, function(x) shapiro.test(x)$p.value)
fligner.test(pVOT ~ Class, data =  syl1)
syl1.1way <-data.Condensed.1way[data.Condensed.1way$Syllables == '1', ]
oneway_test(pVOT ~ Class, data = syl1.1way)
npar.syl.pVOT <- nparcomp(pVOT ~ Class, data = syl1.1way, type = "Tukey")
npar.syl.pVOT$ Analysis

aggregate(Release_ms ~ Class, data = syl1, function(x) shapiro.test(x)$p.value)
fligner.test(Release_ms ~ Class, data =  syl1)
oneway_test(Release_ms ~ Class, data = syl1.1way)
npar.syl.VOT <- nparcomp(Release_ms ~ Class, data = syl1.1way, type = "Tukey")
npar.syl.VOT$ Analysis

# ########################################3
# 1 syllables p/ph and t/th separately
# ########################################3
# P
syl1 <-data.Condensed.1way[data.Condensed.1way$Syllables == '1', ]
syl1.p <-syl1[syl1$Plosive.Letter == 'p', ]
syl1.p.A1 <-syl1.p[syl1.p$Aspiration == '1', ]
tapply(syl1.p.A1$pVOT, syl1.p.A1$Class, mean)
tapply(syl1.p.A1$pVOT, syl1.p.A1$Class, sd)
tapply(syl1.p.A1$pVOT, syl1.p.A1$Class, median)
tapply(syl1.p.A1$pVOT, syl1.p.A1$Class, mad)
aggregate(pVOT ~ Class, data = syl1.p.A1, function(x) shapiro.test(x)$p.value)
syl1.p.A1.mono <- syl1.p.A1[syl1.p.A1$Class == '1', ]
qqPlot(syl1.p.A1.mono$pVOT,
       main="pVOT [ph] mono",
       ylab="")
ggdensity(syl1.p.A1.mono$pVOT,
          main = "pVOT [ph] mono",
          xlab = "pVOT")
syl1.p.A1.multi <- syl1.p.A1[syl1.p.A1$Class == '2', ]
qqPlot(syl1.p.A1.multi$pVOT,
       main="pVOT [ph] multi",
       ylab="")
ggdensity(syl1.p.A1.multi$pVOT,
          main = "pVOT [ph] multi",
          xlab = "pVOT")
syl1.p.A1.o60 <- syl1.p.A1[syl1.p.A1$Class == '3', ]
qqPlot(syl1.p.A1.o60$pVOT,
       main="pVOT [ph] o60",
       ylab="")
ggdensity(syl1.p.A1.o60$pVOT,
          main = "pVOT [ph] o60",
          xlab = "pVOT")

bartlett.test(pVOT ~ Class, data = syl1.p.A1)
pA1.aov <- aov(pVOT ~ Class, data = syl1.p.A1)
summary(pA1.aov)
TukeyHSD(pA1.aov)

# ---------------------------------------------Release VOT
tapply(syl1.p.A1$Release_ms, syl1.p.A1$Class, mean)
tapply(syl1.p.A1$Release_ms, syl1.p.A1$Class, sd)
tapply(syl1.p.A1$Release_ms, syl1.p.A1$Class, median)
tapply(syl1.p.A1$Release_ms, syl1.p.A1$Class, mad)
```

```
aggregate(Release_ms ~ Class, data = syl1.p.A1, function(x) shapiro.test(x)$p.value)
syl1.p.A1.mono <- syl1.p.A1[syl1.p.A1$Class == '1', ]
qqPlot(syl1.p.A1.mono$Release_ms,
       main="VOT [ph] mono",
       ylab="")
ggdensity(syl1.p.A1.mono$Release_ms,
          main = "VOT [ph] mono",
          xlab = "VOT")
syl1.p.A1.multi <- syl1.p.A1[syl1.p.A1$Class == '2', ]
qqPlot(syl1.p.A1.multi$Release_ms,
       main="VOT [ph] multi",
       ylab="")
ggdensity(syl1.p.A1.multi$Release_ms,
          main = "VOT [ph] multi",
          xlab = "VOT")
syl1.p.A1.o60 <- syl1.p.A1[syl1.p.A1$Class == '3', ]
qqPlot(syl1.p.A1.o60$Release_ms,
       main="VOT [ph] o60",
       ylab="")
ggdensity(syl1.p.A1.o60$Release_ms,
          main = "VOT [ph] o60",
          xlab = "VOT")
fligner.test(Release_ms ~ Class, data =  syl1.p.A1)
kruskal.test(Release_ms ~ Class, data = syl1.p.A1)
syl1 <-data[data$Syllables == '1', ]
syl1P <- syl1[syl1$Plosive.Letter == 'p',]
syl1PA1 <- syl1P[syl1P$Aspiration == '1',]
npar.syl1PA1 <- nparcomp(Release_ms ~ Class, data = syl1PA1, type = "Tukey")
npar.syl1PA1$ Analysis
plot(npar.syl1PA1)
###--------------------------------------
# T
syl1.t <-syl1[syl1$Plosive.Letter == 't', ]
syl1.t.A1 <-syl1.t[syl1.t$Aspiration == '1', ]
syl1.t.A0 <-syl1.t[syl1.t$Aspiration == '0', ]
tapply(syl1.t.A0$pVOT, syl1.t.A0$Class, mean)
tapply(syl1.t.A0$pVOT, syl1.t.A0$Class, sd)
tapply(syl1.t.A0$pVOT, syl1.t.A0$Class, median)
tapply(syl1.t.A0$pVOT, syl1.t.A0$Class, mad)
aggregate(pVOT ~ Class, data = syl1.t.A0, function(x) shapiro.test(x)$p.value)
syl1.t.A0.mono <- syl1.t.A0[syl1.t.A0$Class == '1', ]
qqPlot(syl1.t.A0.mono$pVOT,
       main="pVOT [t] mono",
       ylab="")
ggdensity(syl1.t.A0.mono$pVOT,
          main = "pVOT [t] mono",
          xlab = "pVOT")
syl1.t.A0.multi <- syl1.t.A0[syl1.t.A0$Class == '2', ]
qqPlot(syl1.t.A0.multi$pVOT,
       main="pVOT [t] multi",
       ylab="")
ggdensity(syl1.t.A0.multi$pVOT,
          main = "pVOT [t] multi",
          xlab = "pVOT")
syl1.t.A0.o60 <- syl1.t.A0[syl1.t.A0$Class == '3', ]
qqPlot(syl1.t.A0.o60$pVOT,
       main="pVOT [t] o60",
       ylab="")
ggdensity(syl1.t.A0.o60$pVOT,
          main = "pVOT [t] o60",
          xlab = "pVOT")
fligner.test(pVOT ~ Class, data =  syl1.t.A0)
syl1 <-data.Condensed.1way[data.Condensed.1way$Syllables == '1', ]
syl1T <- syl1[syl1$Plosive.Letter=="t",]
syl1TA0.1way <- syl1T[syl1T$Aspiration=="0",]
oneway_test(pVOT ~ Class, data = syl1TA0.1way)
npar.syl.t.A0.pVOT <- nparcomp(pVOT ~ Class, data = syl1.t.A0, type = "Tukey",
   alternative = "two.sided")
npar.syl.t.A0.pVOT$ Analysis
plot(npar.syl.t.A0.pVOT)
###-------------------------------------------------------------
# Release VOT T A0
tapply(syl1.t.A0$Release_ms, syl1.t.A0$Class, mean)
tapply(syl1.t.A0$Release_ms, syl1.t.A0$Class, sd)
tapply(syl1.t.A0$Release_ms, syl1.t.A0$Class, median)
tapply(syl1.t.A0$Release_ms, syl1.t.A0$Class, mad)
aggregate(Release_ms ~ Class, data = syl1.t.A0, function(x) shapiro.test(x)$p.value)
syl1.t.A0.mono <- syl1.t.A0[syl1.t.A0$Class == '1', ]
qqPlot(syl1.t.A0.mono$Release_ms,
       main="VOT [t] mono",
       ylab="")
```

```
ggdensity(syl1.t.A0.mono$Release_ms,
          main = "VOT [t] mono",
          xlab = "VOT")
syl1.t.A0.multi <- syl1.t.A0[syl1.t.A0$Class == '2', ]
qqPlot(syl1.t.A0.multi$Release_ms,
       main="VOT [t] multi",
       ylab="")
ggdensity(syl1.t.A0.multi$Release_ms,
          main = "VOT [t] multi",
          xlab = "VOT")
syl1.t.A0.o60 <- syl1.t.A0[syl1.t.A0$Class == '3', ]
qqPlot(syl1.t.A0.o60$Release_ms,
       main="VOT [t] o60",
       ylab="")
ggdensity(syl1.t.A0.o60$Release_ms,
          main = "VOT [t] o60",
          xlab = "VOT")
fligner.test(Release_ms ~ Class, data =  syl1.t.A0)
oneway_test(Release_ms ~ Class, data = syl1TA0.1way)
npar.syl.t.A0.VOT <- nparcomp(Release_ms ~ Class, data = syl1.t.A0, type = "Tukey")
npar.syl.t.A0.VOT$ Analysis

# ########################################
# Norm by 2 syllables [t]
# ########################################
syl2 <-dataSYL[dataSYL$Syllables == '2', ]
syl2.t <- syl2[syl2$Plosive.Letter == "t",]
syl2.t.A0 <- syl2.t[syl2.t$Aspiration =="0",]
# VOT
tapply(syl2.t.A0$Release_ms, syl2.t.A0$Class, mean)
tapply(syl2.t.A0$Release_ms, syl2.t.A0$Class, sd)
tapply(syl2.t.A0$Release_ms, syl2.t.A0$Class, median)
tapply(syl2.t.A0$Release_ms, syl2.t.A0$Class, mad)
tapply(syl2.t.A0$Release_ms, syl2.t.A0$Class, mean)
aggregate(Release_ms ~ Class, data = syl2.t.A0, function(x) shapiro.test(x)$p.value)
fligner.test(Release_ms ~ Class, data =  syl2.t.A0)
syl2 <-data.Condensed.1way[data.Condensed.1way$Syllables == '2', ]
syl2T <- syl2[syl2$Plosive.Letter=="t",]
syl2TA0.1way <- syl2T[syl2T$Aspiration=="0",]
oneway_test(Release_ms ~ Class, data = syl2TA0.1way)
npar.syl2.t.A0.VOT <- nparcomp(Release_ms ~ Class, data = syl2.t.A0, type = "Tukey")
npar.syl2.t.A0.VOT$ Analysis

#------pVOT
tapply(syl2.t.A0$pVOT, syl2.t.A0$Class, mean)
tapply(syl2.t.A0$pVOT, syl2.t.A0$Class, sd)
tapply(syl2.t.A0$pVOT, syl2.t.A0$Class, median)
tapply(syl2.t.A0$pVOT, syl2.t.A0$Class, mad)
aggregate(pVOT ~ Class, data = syl2.t.A0, function(x) shapiro.test(x)$p.value)
fligner.test(pVOT ~ Class, data =  syl2.t.A0)
oneway_test(pVOT ~ Class, data = syl2TA0.1way)
npar.syl2.t.A0.pVOT <- nparcomp(pVOT ~ Class, data = syl2.t.A0, type = "Tukey")
npar.syl2.t.A0.pVOT$ Analysis
```