Sociophonetic Variation in Afrikaans Vowel Production

Hannes Essfors

u12249855@univie.ac.at

December 16, 2024, Vienna

Introduction

Afrikaans, one of the many official languages of South Africa and, next to English, the only indigenous Indo-European language of Africa, is quite a peculiar language. Although not indigenous to Africa, it was brought to the continent with Dutch settlers who settled at the Cape of Good Hope in 1652 and developed from their vernacular into a distinct language (Britannica, 2023). Originally spoken by white colonists, the union of Whites, native Khoekhoe, and African and Asian slaves resulted in an ethnically mixed social category, "coloured," growing up with Afrikaans as a native language (Britannica, 2024b). Today, the Coloured make up the majority of the Afrikaans speaker community (50.52%), with White constituting 39.78% and the rest of the speakers being Black African (8.84%) and Indian Asian (0.86 %) (Wissing, 2020). The clear social and legal separation of "blacks," "coloured," and "whites" through the apartheid system in the second half of the twentieth century forced the different groups of native Afrikaans speakers to occupy different social spaces (Britannica, 2024a). Due to this separation, the groups enjoyed different access to education and contact with different languages, which would have given rise to different Sociophoentic variations of Afrikaans.

In 2023, Wissing *et al.* carried out an investigation of the vowel spaces of coloured Afrikaans speakers from diffrerent parts of South Africa using word lists and compared the differences between the variants in formant frequencies (F1 and F2) using a method based on overlapping distance (Pastore & Calcagnì, 2019). In this paper, the results of Wissing *et al.* (2023) are built upon and extended by comparing the vowel spaces between that of coloured Afrikaans speakers and white Afrikaans speakers. Furthermore, the methodological aspect of retrieving such recordings is considered by comparing the results based on word lists and readings of texts. Since collecting data using speech in context is cumbersome and resource intensive, understanding the impact on results in the context of Afrikaans would provide valuable insights for future studies.

The aim of the project is to address the following questions:

1. Do coloured and white sociolinguistic variants of Afrikaans differ from each other, how, and to what extent?

2. To what extent do results collected using wordlists contra speech in context match?

3. How can differences between two vowel spaces quantitatively be compared with regard to: effect size, dispersion and contribution of single vowels?

1 Data

The data for the analysis were collected by Daan Wissing (North West University, Potchefstroom, South Africa) and consists of F1 and F2 values for four different groups, here denoted and refered to as *A, B, C* and *D*

1. A: Coloured Afrikaans, reading

2. B: White Afrikaans, reading

3. C: Coloured Afrikaans, wordlist

4. D: White Afrikaans, wordlist

1.1 Data ingestion and processing

The original data was supplied as XLS-files, with one table for each group. The original files were read into R R Core Team (2023) which was used for all processing and analysis. Each data frame contained 6 variables with each record corresponding to a recording of a vowel:

1. Group (white or coloured)
2. Name
3. Vowel
4. Word
5. F1
6. F2

The name variable also contained metadata about the speakers, which, using regular expressions, were attempted to be extracted to potentially enhance the analysis using regular. In the case of A (404 speakers), all genders could be extracted, but information about age was missing for 69 speakers. B (87 speakers) was completely lacking metadata. In the case of C (336 speakers) the gender of all but 3 speakers was deduced, while 66 speakers were missing information about age. For D (9 speakers), the gender of all speakers was concluded to be female based on the names given in the name-column, while no information regarding age was available. It was concluded that doublets of speakers with the same name were present in C, which thus was subjected to extensive cleaning. All strings were removed of its first letter, any "_", and the words "lys", "wordelys", since these provided inconsistensies.

Using the IPA table (Figure 14) found in the Appendix, made available by Andreas Baumann and also used by Wissing et al. (2023), the vowels were remapped to the corresponding IPA orthography. The vowels present in the datasets were /ə/, /ɑ/, /a/, /ɛ/, /e/, /əi/ /ø/, /i/, /ɔ/, /u/, /œ/, /œy/, /o/, /œu/, /y/ with B, C and D covering all vowels but A missing /œİn this process, inconsistencies regarding the encoding of the vowels /ə/ and ø were found in A and B, which consequently were corrected. The text used for the recordings in A and B were equivalent to those of the *Northwind and the Sun*, an Afrikaans version from

Wissing (2020) is found in the Appendix. The Northwind and the Sun is regularly used in phonetics/phonology to sample phoneme inventories Baird et al. (2022), but generally they do not cover the full inventory of languages, which could be why some additional words were found in A and B which are not part of the Northwind and the Sun (uiteindelik, stywer, toegehul, se, later, jas, dan, vas, uittrek, beurt, doen, knoop). However, only removing beurt would have reduced the vowel space, which is why all additional words except beurt were removed. The words used in C and B to sample the vowels using wordlists were: kis, bas baas ses bees byl reus kies bos koek bus, uit, boot oud nuus.

1.2 Data Quality

All datasets were checked for missing values, and 11 datapoints in A, all concerning /ɔ/, with missing F1 and F2 values were identified and removed. As the analysis will take place on a speaker level, the number of vowels covered by each speaker was identified by aggregating the speaker column and counting the unique occurrences of vowels. In A, 300 speakers covered 11 vowels and 87 covered only 3, meaning that no single speaker covered the entire vowel space; in B, 61 speakers covered all 15 vowels; in C, 166 speakers covered all 15 vowels, while in D, 8 of the 9 speakers covered the entire vowel space. To ensure the highest possible data quality, all speakers not covering at least 14 vowels were removed, in the case of A, all speakers covering 11 and 3 vowels respectively were, since the 3 vowels covered by 87 speakers were not covered by any other speakers. A table for each dataset with number of speakers per vowel, instances and mean F1 and F2 values for each dataset can be found in the appendix. From Figures 16, 17, 18 and 19, it is clear that speakers have multiple recordings for each vowel, which makes those data points dependent. To avoid this and reduce intraspeaker variability, the median F1 and F2 for each speaker and each vowel was calculated and used instead. That way, there is only one data point for each speaker-vowel-combination. To also reduce interspeaker variability, the formants were normalized using lobanov normalization (Lobanov, 1971) and thus z-transforming all the formant values for each speaker in relationship to the entire distribution of that speaker. As a result, the units of the transformed formants are given in standard deviations instead of Herz.

2 Descriptive statistics

After processing and cleaning the data, summary statistics were calculated and are presented in Table 1. It can be noted that the median F1 and F2 of D is deviating somewhat from that of the other Datasets. A, B and C observe some skew of their F1 values F2 is somewhat skewed for B. It should also be noted that the sizes of the datasets are very unbalanced, with comparable number of instances in A and C (3561 and 3246 respectively), but only 1209 instances for B and the abysmal 134 of D.

Set	n instances	n speakers	n vowels	median F1	median F2	skew F1	skew F2
A	3561	387	14	-0.09	-0.04	0.58	0.10
B	1209	82	15	-0.01	-0.02	0.10	0.28
C	3246	220	15	-0.12	0.01	0.64	-0.10
D	134	9	15	-0.30	-0.12	0.78	0.10

Table 1: Summary statistics for all datasets. Formant values are given in standard deviations

Looking at the distribution of Formants for the datasets in Figure 1, it can be observed that the distributions are multimodal, and that the distributions of C and D are somewhat similar, especially with regard to F1, while b and A show some similarity with regard to F2.

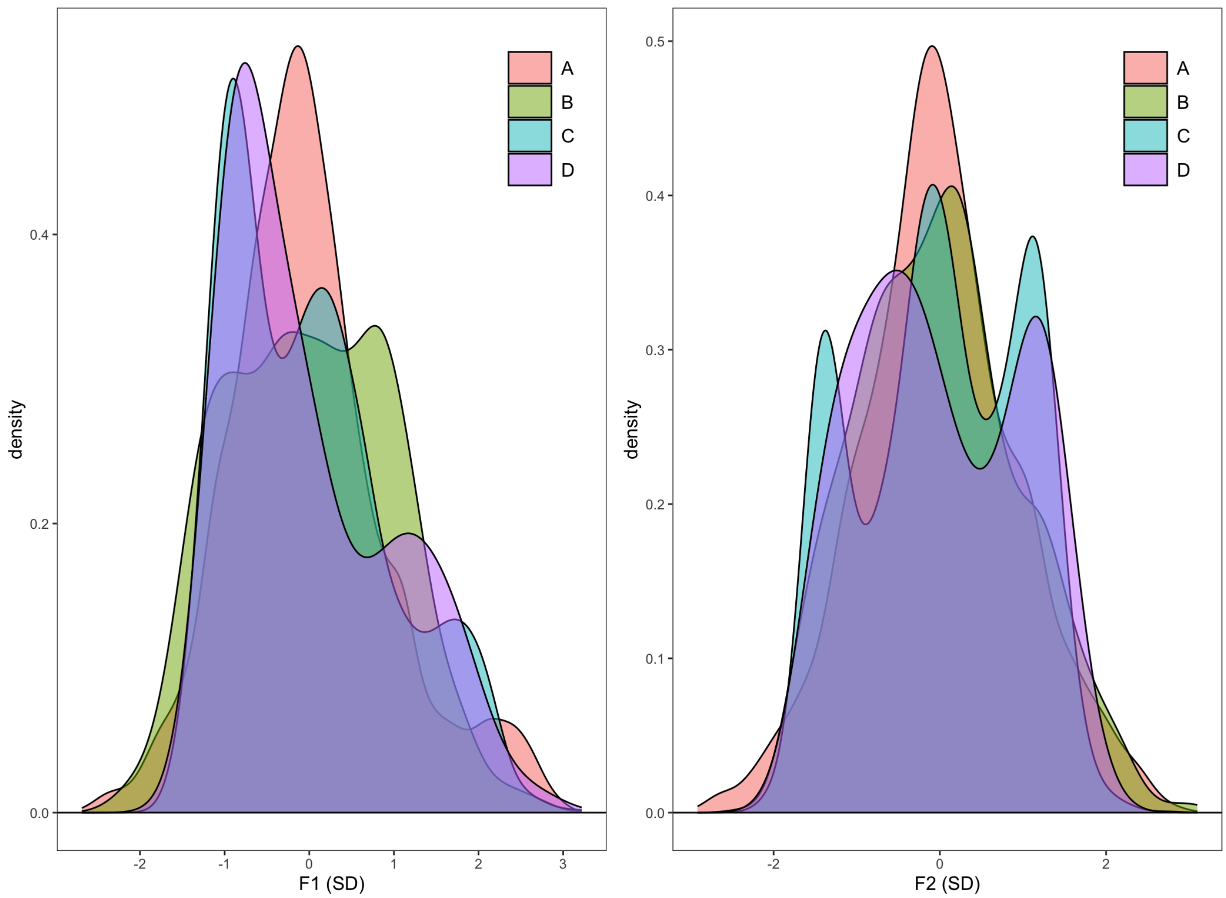

Figure 1: Distribution of F1 and F2 values for all datasets

The distribution per vowel can be further visualized using boxplots as seen in Figure 2. Studying the boxplots, it appears that C, B and D follow each other fairly closely, while the distributions of A are deviating quite clearly. The most evident example of this are found for the F2 values of /ø/ and /y/ where B, C and D have positive values while the corresponding A values are negative.

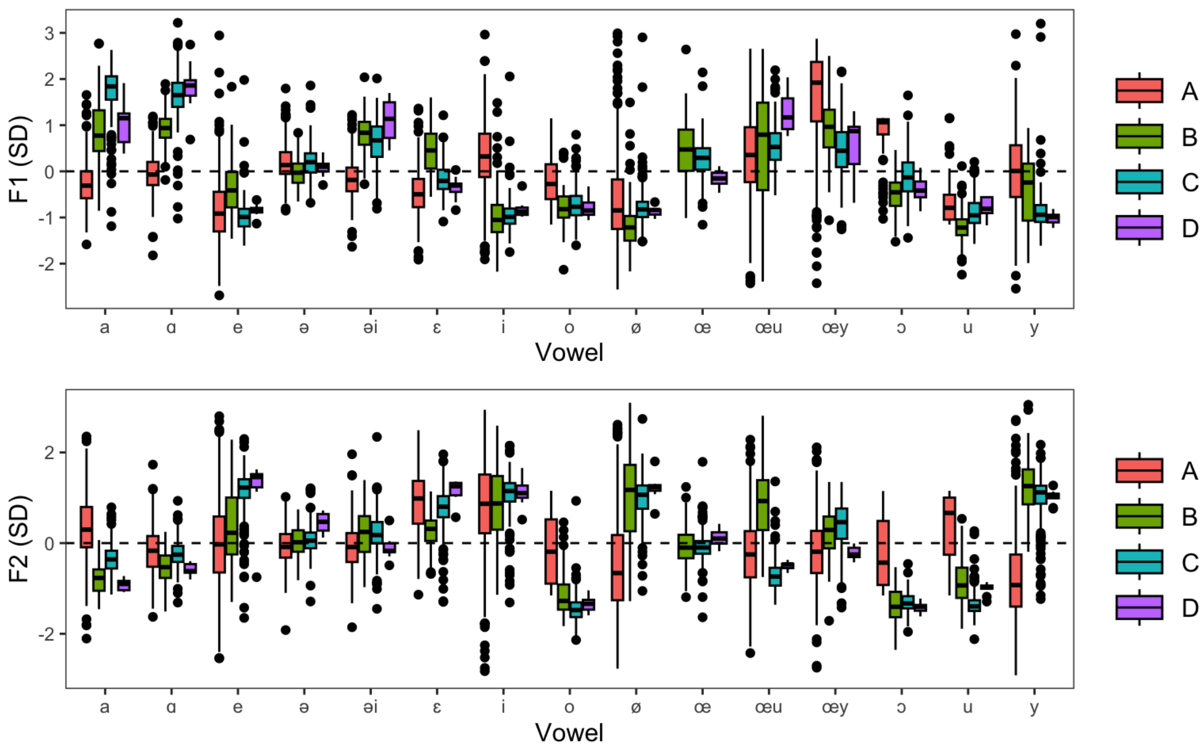

Figure 2: Boxplots of F1 and F2 values for all datasets per vowel

Further visualizing the vowel spaces, F1 vs F2 plots (with inverted x- and y-axis') were plotted for all datasets and can be seen in figure 3. Clear vowel triangle can be seen for C and D, with very clearly defined vowels. In A and B there is quite a lot of overlap between the vowels, signalling that there is large within vowel variability. In B the back-, front- and mid-vowels clearly can be distinguished, but in A, they appear to be more or less random. Due to this, A was deemed unsuitable for further analysis and therefore excluded. Since both A and B are datasets recorded using speech in context, it could, with regard to research question (1), be suggested that speech in context recordings are associated with increased intravowel variability, but it can't definitely be concluded. These results however suggests that using wordlists to sample formants provide more reliable results in the context of research on Afrikaans.

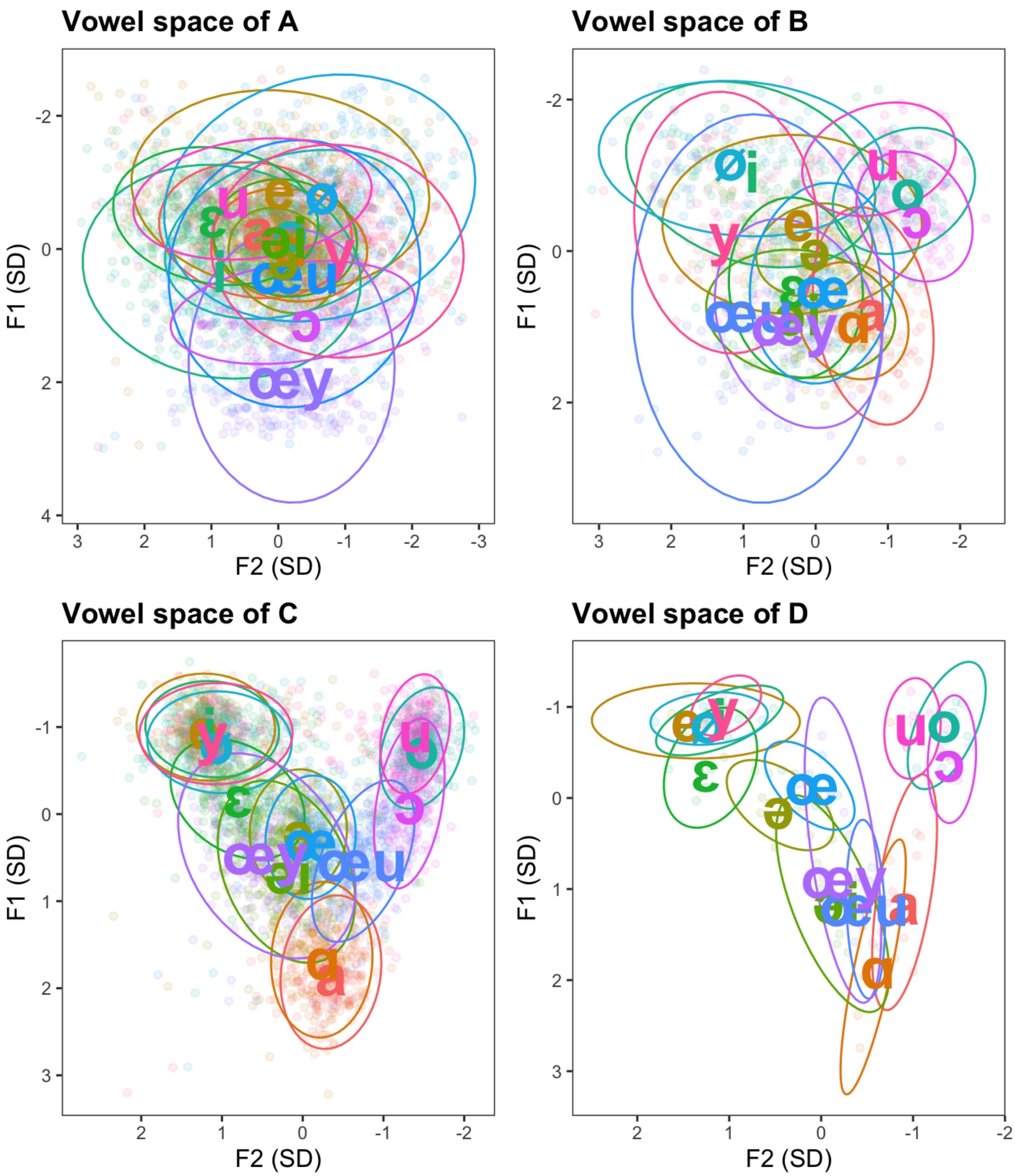

Figure 3: Vowel spaces of all datasets with inverted axis'. Ellipses mark 95% confidence regions for each vowel

3 Analysis

3.1 Approaches

To compare the differences between the vowel spaces, three approaches were used and compared:

3.1.1 Overlapping approach

The distribution overlapping approach introduced by Pastore & Calcagnì (2019) is implemented in the same way as in Wissing *et al.* (2023, eq (1)): the overlapping coefficient η (0 $\rightarrow$ no overlap; 1 $\rightarrow$ identical distributions) is calculated for the F1 and F2 distributions of each vowel in both vowel spaces respectively and the complement δ = 1 - η is used as a distance measure. An index D for two vowel spaces p and q is then defined as

$$D(p,q) = \frac{1}{2 \cdot |V|} \cdot \sum_{v \in V} (\delta(F1_{v,p}, F1_{v,q}) + \delta(F2_{v,p}, F2_{v,q})) \tag{1}$$

Where V is the set of vowels v constituting the vowel spaces. If D(p,q) = 1 it implies that none of the vowels in the vowel spaces p and q overlap in neither F1 nor F2, while D(p,q) = 0 implies that all vowels have identical distributions in both vowel spaces. This index will be referred to as the overlap index.

This approach has the strength of being non-parametric, i.e. it does not assume a given distribution of the data; it is restricted to values between 0 and 1; and in the context of this project it allows direct comparison with the results of previous research on Afrikaans employing the same method (Wissing et al., 2023). It also has an intuitive interpretation which is easy to visualize by simply plotting two distributions over each other.

The weakness of the approach lies in its lack of directionality and only non-explicit sensitivity to difference in central tendency: δ = 0.50 implies that 50 percent of the areas of the two distributions are shared, but does not imply whether one of the distributions tends to larger or smaller values, e.g. two distributions could have the same mean F1 but different variances, resulting in a large overlapping distance, even tough the central tendency is equal. If indeed one of the distributions are shifted to smaller or larger values, which one of the distributions are shifted or tends to larger values is not explicitly given by the measure, since it gives the absolute difference. Furthermore, the overlapping distance is potentially insensitive to outliers, since extreme values at both tails of the distributions could disproportionally elongate the distribution thus inflating the distance if only present in one distribution.

3.1.2 Domination approach

The weakness of the overlapping distance as described above prompted a novel approach using Cliff's 1993 delta. Cliff's delta is a non-parametric effect size for comparing the overlap of the distributions, similarly to overlap distance, although not as direct in the sense of expressing the difference of shared area of the distributions. It instead indicates the percentage of values in one distribution that are larger or smaller than the values in the other distribution. For two groups, x and y, this is formalized as

$$Cliff's\ \delta = \frac{\#\,(x > y) - \#\,(x < y)}{n_x \cdot n_y} \tag{2}$$

where the number of instances where values of x are smaller than values of y are subtracted from the number of instances where the values of x are larger than the values of y (Meisel & Yao, 2024, eq (1)). The result is then normalized by the total number of instances of x and y, where Cliff's $\delta = 0$ implies that there are equally many cases of x being larger than y as cases of x being smaller than y, i.e. the distributions are entirely overlapping. Cliff's $\delta = 1$ would imply that all values in x are larger than all values in y i.e. the distributions are not overlapping and x is shifted to larger values, while cliff's δ = -1 indicates no overlap and x is shifted to smaller values relative to y. This solves the problem of directionality: Cliffs delta explicitly indicates if the distribution of a vowel v tends to have larger or smaller F1 values in the vowel space of p compared to the vowel space of q. In such a context, equation (2) can be generalized as

$$Cliff's\ \delta(v_{p,q}) = \frac{\#\,(F1_{v_p} > F1_{v_q}) - \#\,(F1_{v_p} < F1_{v_q})}{n_{v_p} \cdot n_{v_q}} \tag{3}$$

where v is a vowel existing in both p and q. *Mutatis mutandis*, the same equation can be applied to F2 as well.

To compare entire vowel spaces, an index can be derived in the same way as in equation (1), but using the absolute value of Cliff's δ to offset the directionality such that

$$Cliff's\ D(p,q) = \frac{1}{2 \cdot |V|} \cdot \sum_{v \in V} (|Cliff's\ \delta(F1_{v,p}, F1_{v,q})| + |Cliff's\ \delta(F2_{v,p}, F2_{v,q})|) \quad (4)$$

and will be referred to as Cliff's index.

3.1.3 Spatial approach

Since the lobanov normalized F1 and F2 formants together span a two-dimensional space, a spatial approach to measure the literal distance between the vowel spaces as a whole and their components (the vowels) is an intuitive and easy to interpret approach to determining effect size. For this purpose, a centroid C for each vowel v in the respective vowel space is calculated using the median of the F1 and F2 distributions of v as coordinates, such that

$$C_v = (med(F2_v), med(F1_v)). \quad (5)$$

And it follows that a vowel space S is defined as a set of such centroids where

$$S = \{C_v | v \in V\} \quad (6)$$

which allows the structural representation of the vowel space as a convex hull enclosing these points:

$$Conv(S) = \{\lambda_1 C_{v_1} + \lambda_2 C_{v_2} + \cdots + \lambda_k C_{v_k} \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1\} \quad (7)$$

where $k = |V|$. The subset of centroids spanning the convex hull is derived using the algorithm of (Eddy, 1977) and the smallest possible convex hull for those points are derived using the quickhull algorithm (Barber B. & Huhdanpaa, 1977). By deriving such hulls for both vowels space, the area covered by each hull can be calculated, allowing a comparison of the geometric area covered by the spaces. This was done by comparing the absolute difference in area between the vowel space and by calculating the intersection under union IoU, which is given by the intersection between the two convex hulls divided by their union (Jaccard, 1912). For two vowel spaces s and q, the IoU is defined as

$$IoU(p,q) = \frac{|Conv(p) \cap Conv(q)|}{|Conv(p) \cap Conv(q)|} \tag{8}$$

which can be turned into a distance measure as

$$d_{IoU}(p,q) = 1 - IoU. \tag{9}$$

This functions as an index since $0 \leq d_{IoU} \leq 1$, and can be used to quantify entire vowel spaces. Theoretically, such distances could be calculated for each vowel in the two vowel spaces individually without the centroids and turned into an index as in equation 1, but should then be treated with caution since the measure is hypersensitive to outliers. For the entire vowels space, this is unproblematic, since deriving centroids using medians introduces robustness in this regard.

Another option is to calculate Euclidean distances d_{euclid} between each vowel in the two vowel spaces such that

$$d_{euclid}(v_p, v_q) = \sqrt{(F1_{v_p} - F1_{v_q})^2 + (F1_{v_p} - F1_{v_q})^2}. \tag{10}$$

Using the distance measures for each vowel, an effect size can be calculated for the entire vowel spaces by averaging the distance between the vowels. As this is a distance measure in its most literal sense, it should be noted that the Euclidean distance is not restricted to values between 0 and 1 as is the case of overlapping distance for example.

3.1.4 Contribution

To calculate the contribution of a vowel v vowels to the difference between the spaces, the relative contribution can be derived by taking the distance measurement between for vowel v in the spaces p and q The contribution of a single vowel to the difference between the entire spaces can be derived by dividing the measure For all three of the above measurement approaches, the contribution of each vowel to the measure for the vowel spaces as a whole was quantified by taking the effect sized calculated for the vowel, and dividing it by the sum of the effect size of all other vowel, thus turning it into a score, where the sum of all scores is 1.

3.2 Comparing Overlapping distance and Cliff's Delta

Since overlapping distance and Cliff's Delta both function as non-parametric effect sizes of the difference between distributions, it is appropriate to compare their measures to understand the similarities and differences of their behaviour. This was done in an empirical way by means of simulation to study the suggested differences between the measures as they were discussed in the previous section. Distributions were generated by sampling from normal distributions 500 times with random means and variances. For every round of sampling, two samples (n = 100) were sampled with different means and equal variance (actual values random) and combined into a single bimodal sample (n = 200). In addition, a single sample of size 200 with random mean and random variance was sampled to represent a unimodal distribution. As a result, 500 datapoints each containing one bimodal and one unimodal distribution were generated, each of size 200 with random means and variance. For a pair of distributions, Cliff's delta and overlapping distance were calculated for the two distributions together with the difference in mean, standard deviation and bimodality as measured by the bimodality coefficient as in Pfister Roland (2013). Since Cliffs Delta has a directionality, the absolute value of cliff's delta was used, restricting it to [0, 1] to make the measures comparable. The result of the comparison can be seen in figure 4

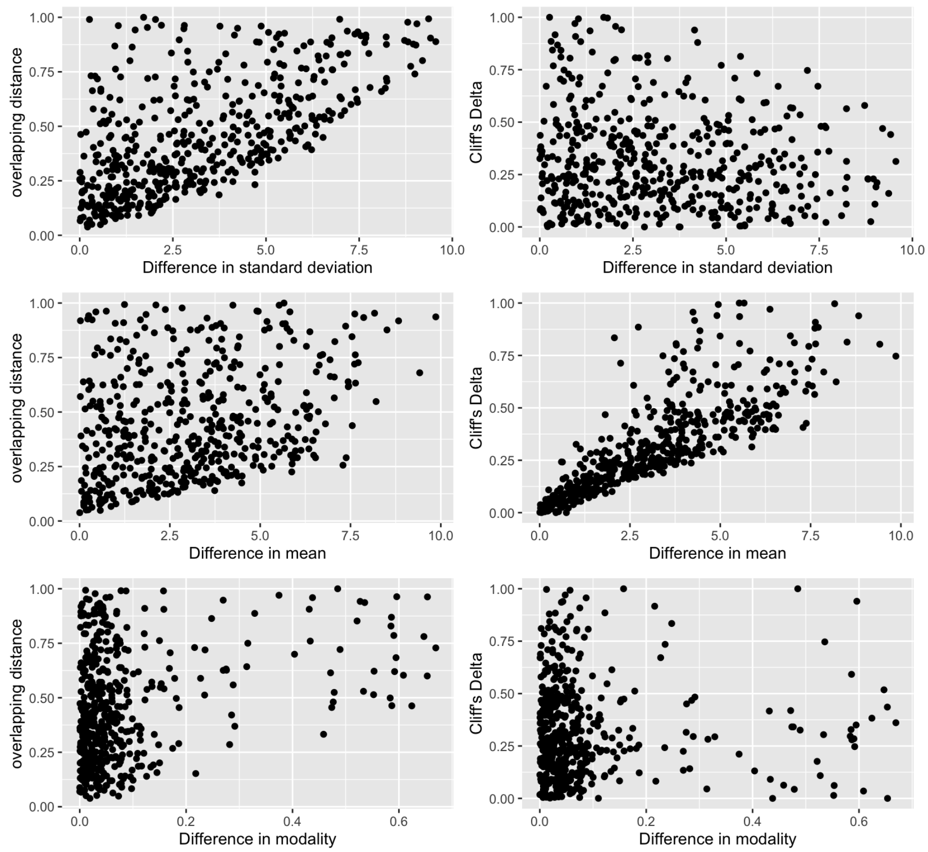

Figure 4: Scatterplots for overlapping distance and cliffs delta compared to differences in mean, standard deviaton and modality

Considering the correlation plots in figure 1, the most evident linear relationships are found between overlapping distance and difference in standard deviation (ρ = 0.62, p < 0.001) and Cliff's Delta and difference in mean (ρ = 0.86, p < 0.001). Overlapping distance and difference in mean see a somewhat weaker positive correlation (ρ = 0.31, p < 0.001), while Cliff's delta and standard deviation show a negligible although significant negative correlation (ρ = -0.08, p < 0.05). Regarding bimodality, there is a weak positive correlation between overlapping distance and difference in bimodality (ρ = 0.22, p < 0.001), but no significant correlation with Cliff's Delta (ρ = 0.020, p = 0.665). This suggests that Cliff's Delta manages to capture the magnitude of the difference while being insensitive to the shape of the distribution. Overlapping distance however mainly captures difference in variance and to some extent modality, but it does not reliably reflect any shifts in the distributions relative to each other. Furthermore, considering the correlation between Cliff's Delta and overlapping distance for all distributions as seen in figure 5 it can be noted that the two measures certainly are related ($\rho = 0.52$, $p < 0.01$), and it even seems that there exists a limit for the maximum value that Cliff's delta can take on for two given distributions, which is given as a function of the overlapping distance between the same distributions.

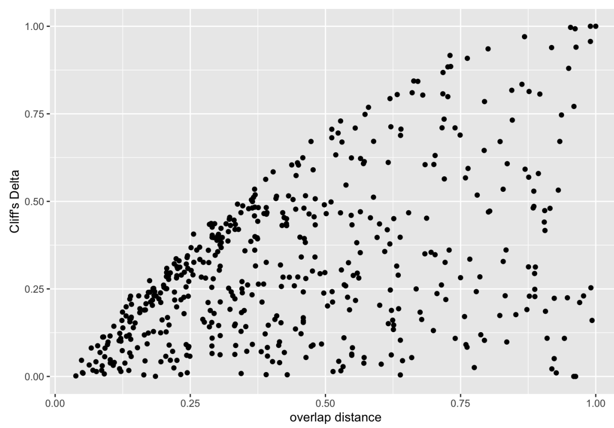

Figure 5: Scatterplot of overlapping distance and Cliff's delta.

Considering figure 5, it can be postulated that for two given distributions, if the overlapping distance between them is small, it follows that Cliff's delta is small. If Cliff's Delta is large, then it follows that the overlapping distance is large as well.

A more common metric than both overlapping distance and Cliff's Delta is Cohen's d, which also measures the degree of non-overlap, assuming that both distributions have the same variance (Meissel, 2024). Since Cohen's d commonly is interpreted as "small", "moderate" and "large", this can be assessed for Cliff's Delta and overlapping distance by simulating normal distributions with fixed variation and varying means and then calculating all effect sizes for pair wise distributions. The approximated results of such a simulation are found in table 3.2

Interpretation	Cohen's d	Cliff's δ	Overlapping Distance
small	0.20	0.11	0.11
medium	0.50	0.27	0.20
large	0.80	0.43	0.29

Table 2: Conversion of Cohen's d to Cliff's δ and overlapping distance

3.3 C vs D

The first two datasets to analyse were C and D. C consists of Black speakers and D of white speakers, both with the wordlist recording method. Considering figure 3, the vowel spaces are fairly similar. However, some differences can be noted. Firstly, the mid-vowels /ɛ/, /ə/, and /œ/ appear to be somewhat fronted by the white speakers, with the central vowels /ə/ and /œ/ being raised as well. The back mid-vowel /ɔ/ of the white speakers is also raised somewhat, while the back high-vowel /u/ is somewhat lowered and clearly fronted. The most pronounced difference is found for /a/, which is clearly raised and backed by the white speakers.

The diphthongs /əi/, /œy/, and /œu/ are more clearly set apart for the coloured speakers, while they for the white speakers cover more or less the same space which is lowered and backed compared to the coloured speakers. To which extent this actually is an observed difference should be considered with caution, as diphthongs are more complex than monophthongs, and we are forced to assume that the measures were derived in the same way for both C and D, but we are not sure. From the looks of it, it almost seems as if it was calculated using the starting point of the formant instead of taking the average in the case of D.

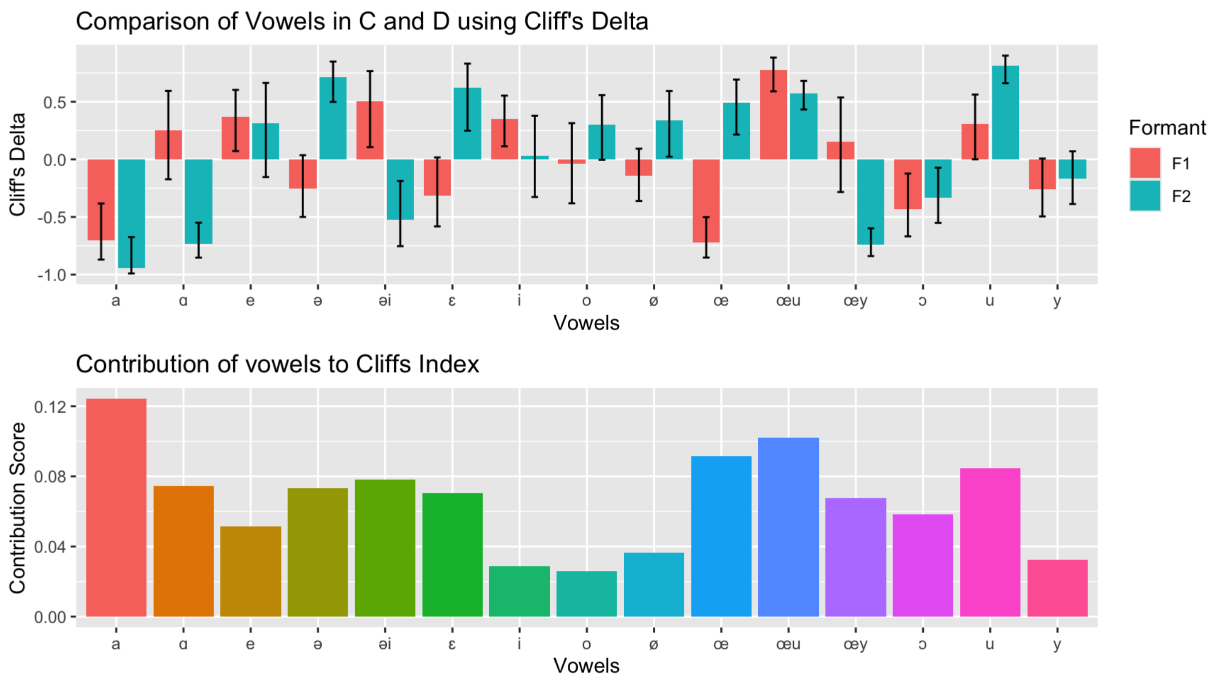

Figure 6: Barplots with Cliff's Delta calculated for F1 and F2 for all vowels in the vowel spaces of C and D with 95 % confidence intervals and the contribution of the difference between the vowels to the index of the vowel space. C is used as a reference.

A more exact comparison of the vowel space is given by considering the distance measures and the contribution of the different vowels. Viewing the Cliff's Delta calculated for the pairs of vowels across C and D with C as a reference as can be seen in figure 6, the visual analysis of the vowel spaces can be confirmed. The largest effect is seen for A, which have shifted F1 and F2 with a negative coefficient, indicating that F1 and F2 values are lower in D relative to D. The results also indicates a significant and large effect for /ɑ/ which is backed in D relative to C. Furthermore, medium to large effects are found for the mid-vowels. Cliff's Delta does not indicate a significant raising of /e/ however, but the fronting effect is large. Similarly, /ɛ/ is not significantly raised but well fronted. For /əi/ on the other hand, a significant fronting and raising is seen, indicating that this is a vowel that clearly sets the coloured and white speakers apart. As noted in the visual presentation, /u/ is fronted in D but not significantly lowered, while /ɔ/ is both lowered and backed. Finally, there are significant differences between the diphthongs as indicated in the vowel space plot, while

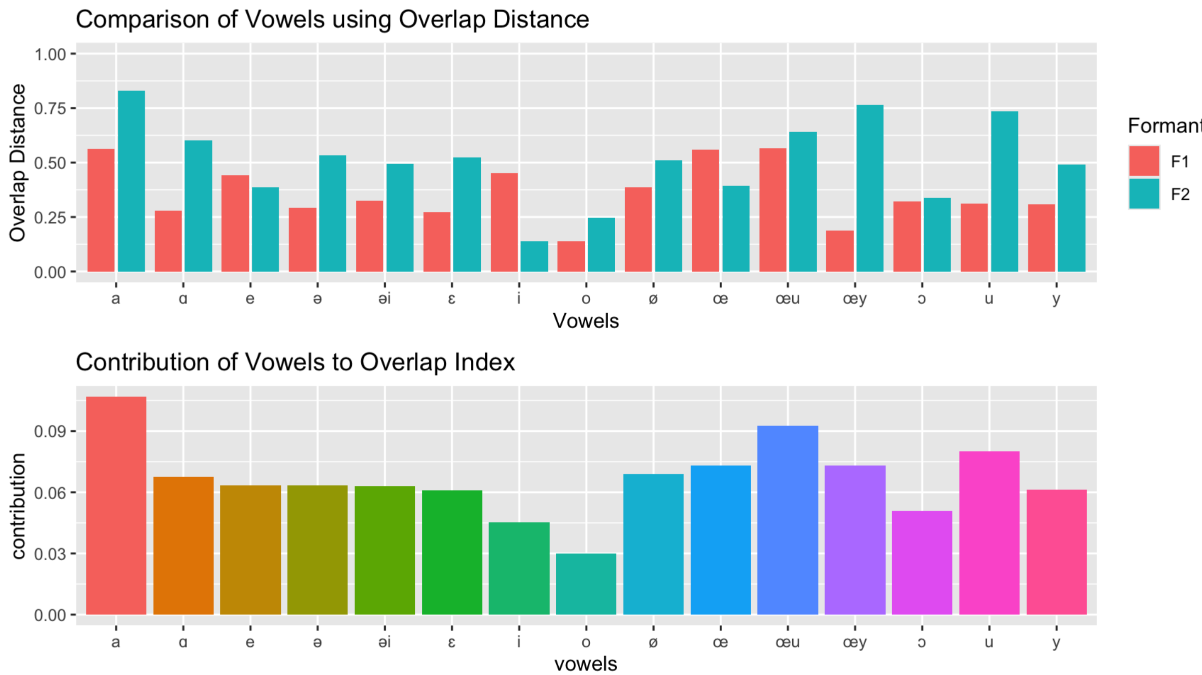

Figure 7: Barplots with overlapping distance calculated for F1 and F2 for all vowels in the vowel spaces of C and D with 95 % confidence intervals and the contribution of the difference between the vowels to the index of the vowel space. C is used as a reference

generally no significant differences are found for the front high-vowels.

Considering the contribution to the index, /a/ is by far the largest contributor, together with /u/ and the mid-vowels, if one does not consider the diphthongs as well. Thus, in conclusion, these results suggest that white speakers pronounce their /a/ phonemes more raised and backed, their /u/ more fronted and their central and front mid-vowels fronted.

To get an even clearer picture, the approach using overlap distance can also be considered, as seen in 7. The same general trend is seen as for Cliff's Delta, with the effect size for /a/ being the largest. Similarly, large effect sizes are found for /u/, the mid-vowels and the diphthongs. Interesting to note is that the contribution of vowels to the overlap index is generally more even, and the front-high vowels contribute more. This would suggest that there are differences in the distribution of the front-high vowels between white and coloured speakers, but not so much in the distinctness of the place of articulation, but rather that between the groups, the variation in how the phoneme is realized is different

Finally, the geometric measures (Figure 8) can give a third perspective on the differ-

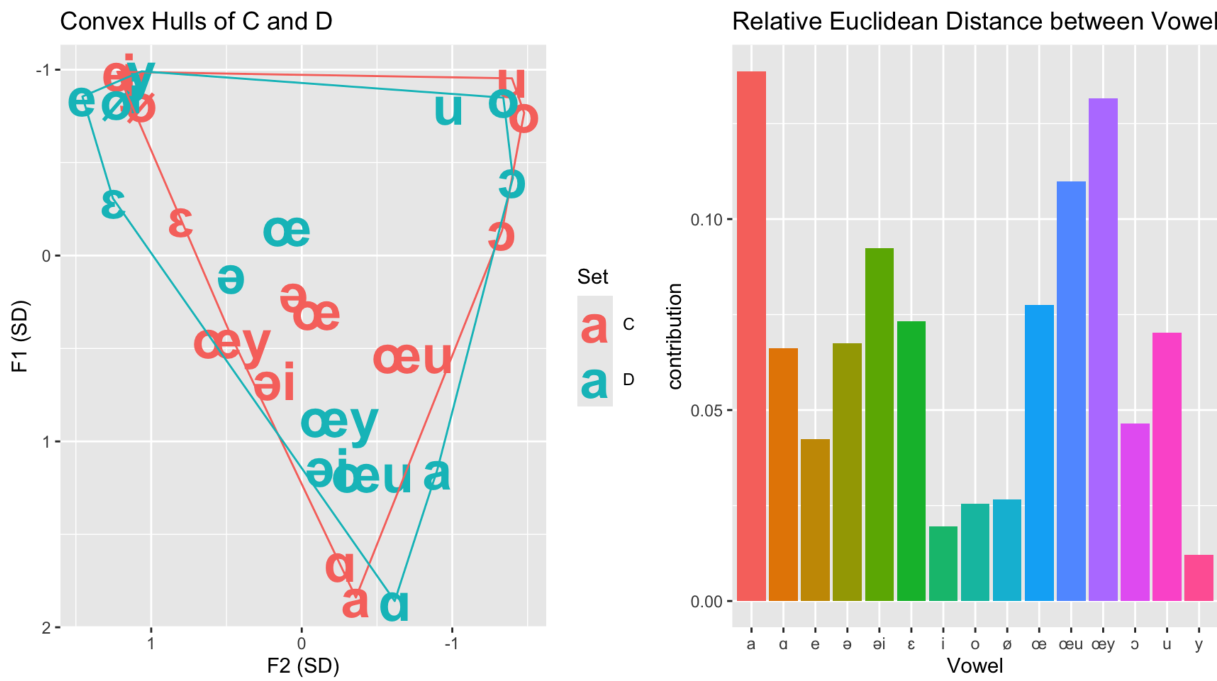

Figure 8: The geometric vowel spaces of C and with their convex hulls plotted over each other and the relative Euclidean distance between each vowel pair.

ences between the distributions. Differences in the convex hulls are found mainly due to the fronting of /e/ and /ɛ/ and the backing of /a/ and /ɑ/ in the vowel space of D. Considering the relative Euclidean distance between vowels, a similar result to that of Cliff's Delta in figure 6 is found, with the clearest difference being inflated contribution of diphthongs.

To summarize the different approaches, a correlation map between the contributions of vowels to the measures can be considered, as seen in Figure 6. A combination of all three contribution graphs can be found in the appendix. Overlapping distance and Euclidean distance are correlated (r = 0.75) which is a good sign, but not it as strong as the correlation between Cliff's Delta and the Euclidean distance (r = 0.86) and between Cliff's Delta and overlapping distance (r = 0.86). This is a strong argument for the strength of using Cliff's delta, since Euclidean distance only take the difference between the centroids (medians) and does not consider the shape of the distribution, it suggests that Cliff's Delta manages to capture the differences in centrality and variance in a way that neither overlapping distance nor Euclidean distance does on their own. This coupled with directionality suggests that Cliff's delta is a well suited measurement for comparing vowel spaces, especially if the dis-

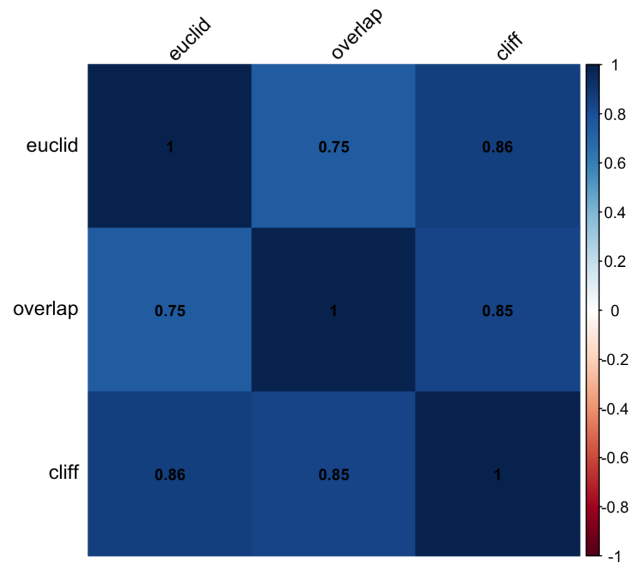

Figure 9: Correlationplot for the indices based on relative contribution of vowels

tinctions of the distribution is of interest. Reporting it together with overlap index would be ideal to understand if the differences are reflecting of a shift in the place of articulation or of variance between the speakers.

3.4 B vs D

B consists of vowel recordings of white speakers reading a text, and D consists of vowel recordings from white speakers reading lists of words. Looking at the vowel spaces in figure 3, large variability in the vowel realizations of the speakers in B is immediately noticed (size of the ellipses). This intuitively makes sense, since reading a text creates complexity, with more factors influencing the pronunciation such as speed, intonation, stress, which could explain an increased variability in how the vowels are realized. In the reading, there were also more unique words for each vowel, which most likely means that single vowels appear in different sound contexts, which might colour the pronunciation of the vowel, thus increasing the variability for B.

Regarding single vowels, a clear lowering of /y/ and /e/ is noticed together with a backing of /e/, realizing it as a mid central vowel. Similarly, /ɛ/ is substantially lowered and backed as well. The mid central vowels /ə/ and /œ/ see some changes as well, with /ə/

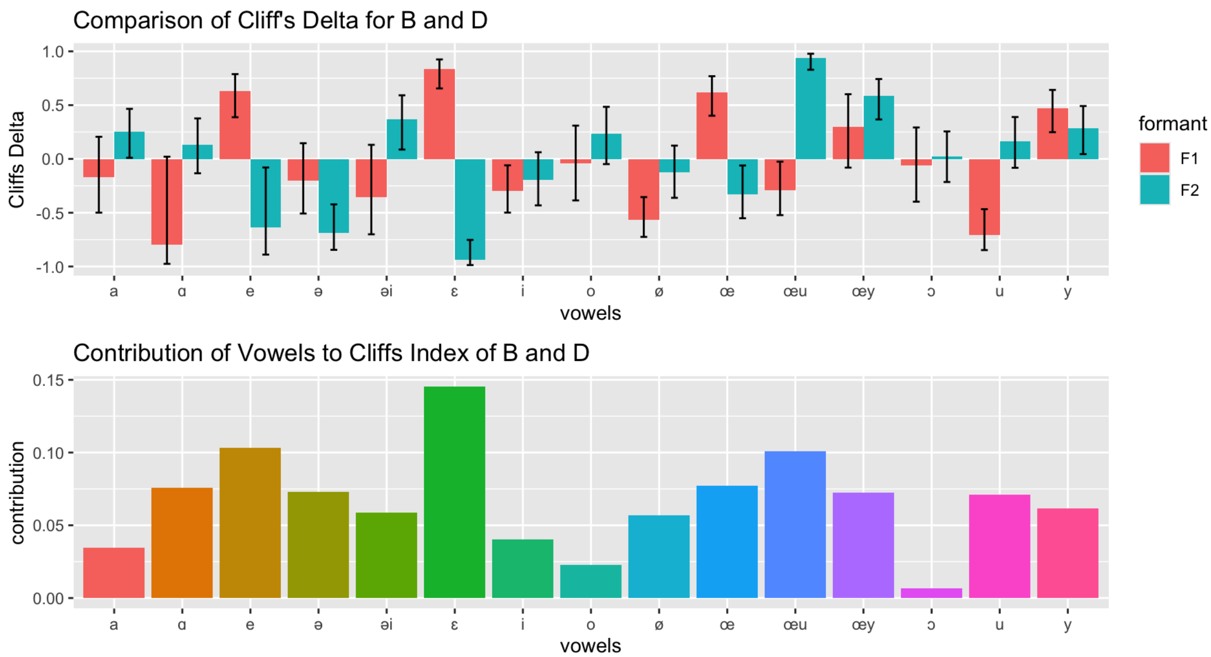

Figure 10: Barplots with Cliff's Delta calculated for F1 and F2 for all vowels in the vowel spaces of B and D with 95 % confidence intervals and the contribution of the difference between the vowels to the index of the vowel space. D is used as a reference.

backed and /œ/ lowered.

The diphthongs are noticeably fronted, with the fronting of /œu/ being especially peculiar and /ɑ/ is clearly raised. Considering /a/ and /ɑ/, they show similar F2s in the vowel spaces, supporting the suggestion of the analysis for C vs D that backed a-sounds is a prominent feature setting white from coloured Afrikaans apart. Similar, these results support that /u/ is somewhat fronted since this is present in both B and D, however in B, /u/ is also raised substantially.

Considering the effect sizes calculated using cliffs delta as seen in figure 10, only few statistically significant changes are found. A raising of /ɑ/ in B can not be supported, and many significant changes are barely significant due to very large confidence intervals. The most clear differences are found for /ɛ/, /e/, /ə/ and the diphthongs, which contribute the most to the index. Generally, this suggests that using a text instead of word lists affect how diphthongs and mid central vowels are realized in the context of research on Afrikaans.

Looking further at the overlapping distance between the vowel distributions in figure 11, a somewhat similar although still remarkably different picture is painted. Again, a more even distribution of contribution as compared to Cliff's Index is observed, and the vowels /e/, /ɛ/ and /œu/ are identified as large contributors, but some vowels seem much more influential such as /y/, /u/ and especially /ø/. This suggests that for these vowels, the difference lies not so much in a shift in place of articulation, but rather a difference in variability of how the vowel is realized.

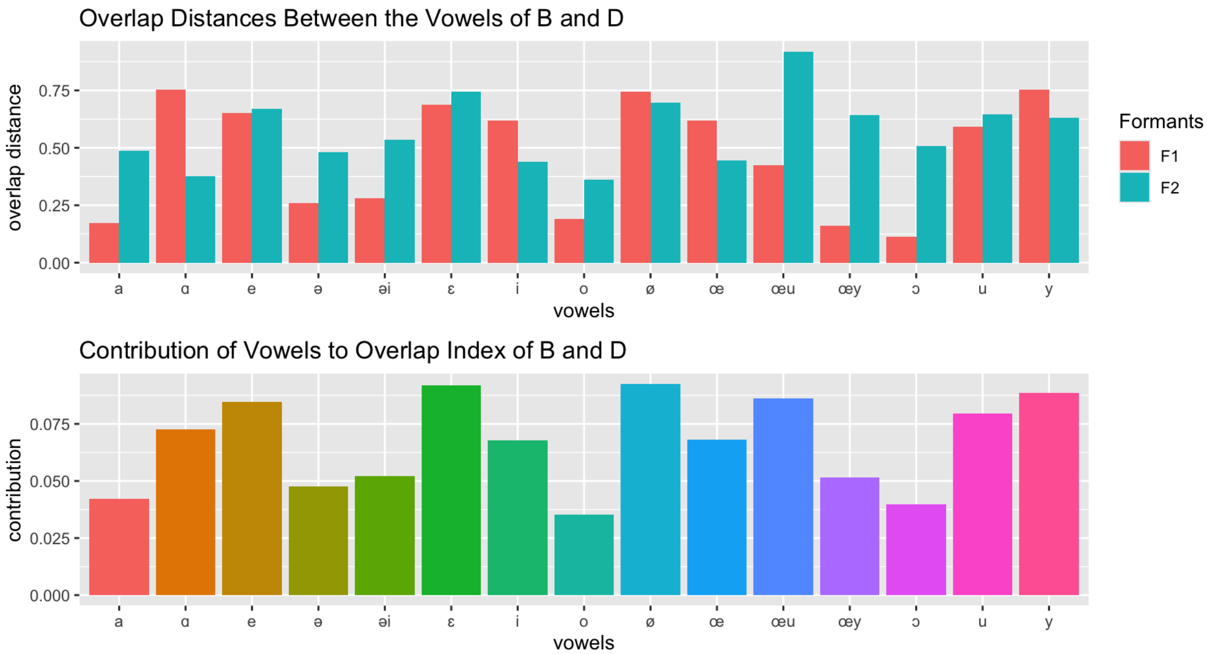

Figure 11: Barplots with Cliff's Delta calculated for F1 and F2 for all vowels in the vowel spaces of B and D with 95 % confidence intervals and the contribution of the difference between the vowels to the index of the vowel space

Moving on to the geometric measures, the convex hulls which are plotted in figure 12, seems to suggest that the entire vowels space of B is raised compared to C. The shape of a convex hulls does however not take shifts of mid central vowels into account, as vowels on the edge of the space can affect its shape. Considering the contribution of vowels to Euclidean distance, again a distribution more akin to the contribution to Cliff's Index is seen. /e/, /ɛ/ and /œu/ are again identified as the most important contributors, but in this case, /œu/ and not /ɛ/ has the largest impact on the difference, which means these two vowels have the most distant centroids.

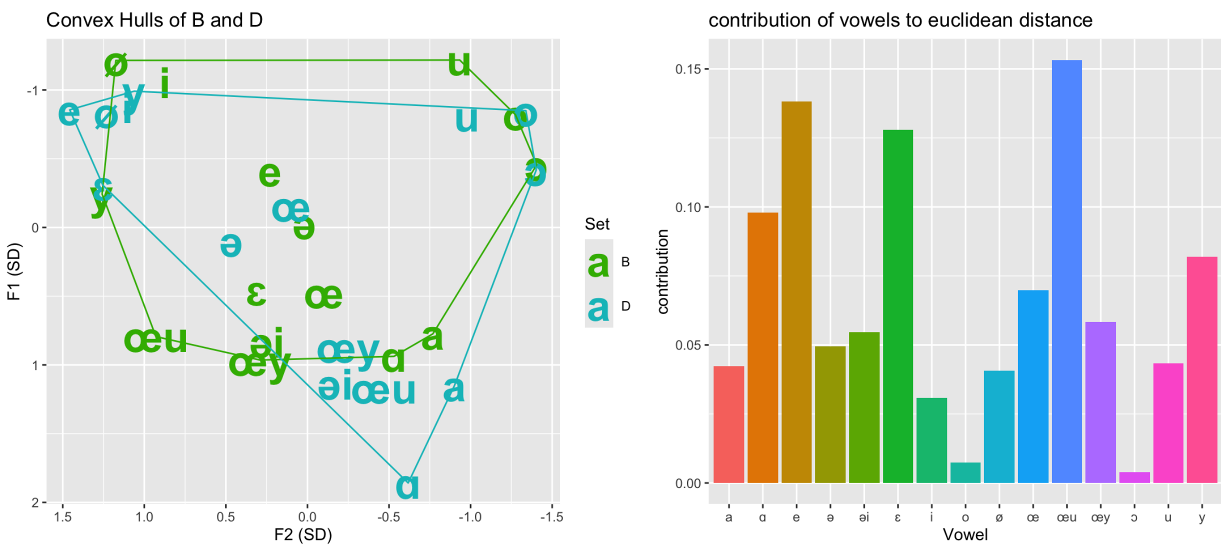

Figure 12: The geometric vowel spaces of B and D with their convex hulls plotted over each other and the relative Euclidean distance between each vowel pair.

Creating a correlation plot to compare the contribution of vowels as seen in figure 13 confirms what was previously concluded, namely that the contributions of vowels are more similar for Euclidean distance and Cliff's Delta (r = 0.87), while overlap, still correlated with the Cliff's Delta (r = 0.69) and Euclidean distance (r = 0.68), seems to also capture some other aspects of the difference between the vowel space. This could be due to a higher variability in how single vowels are realized in B, since it was concluded that overlapping distance is more strongly correlated with difference in variance between distributions than Cliffs delta, and Euclidean distance inherently is insensitive to variance. All in all, this highlights how the measures complement each other.

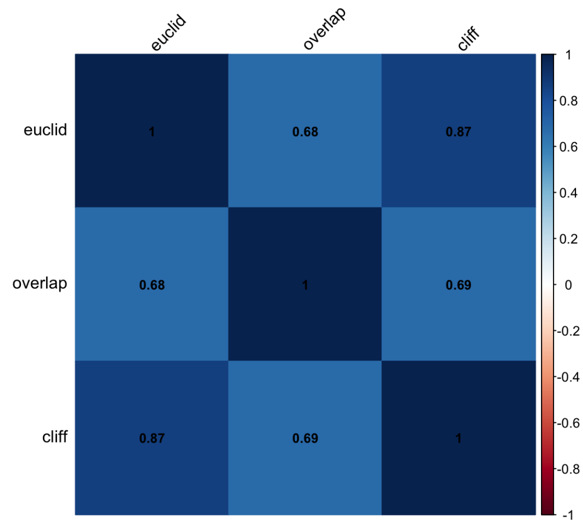

Figure 13: Correlationplot for the indices based on relative contribution of vowels

3.5 Vowel space indices

The comparison of the vowel spaces can be summarized by the composite indices, which are seen here in table 3

Datasets	Cliff's index	Overlap index
C vs D	0.44	0.434
B vs D	0.407	0.52

Table 3: Summary of distance measures for the comparison of vowel spaces C and D, and B and D.

It is interesting to note that the vowels spaces of C and D has a more or less equal Cliff's index and Overlap index, while B and D has a lower Cliff's index and a larger Overlap index. Based on the observed behaviour of the indices, this would suggest that in the vowel space of C and D, the vowels are more displaced on the space relative to each other as indicated by a larger Cliff's index. In B and D, the lower cliff's index indicate a smaller degree of difference in place of articulation of the vowels, while the larger Overlap index indicate that the realizations of vowels in B and D is more varied in between the vowel spaces. Considering the sheer magnitude of the differences between coloured and white speakers (C and D), it is interesting to note an overlap index of 0.434, which can be compared to the difference in overlap index between groups of coloured Afrikaans speakers from different regions observed by Wissing et al. (2023) where the largest observed difference was between the regions of Gobais and South Cape of 0.429, which is roughly the same, although smaller, than that between coloured and white speakers.

4 Conclusion

The vowel spaces of 4 groups of Afrikaans speakers have been compared based on sociophonetic differences (coloured vs white) and vowel sampling technical differences (wordlist vs reading). The results suggest that in the context of Afrikaans phonetic research, using readings to extract samples of vowel realization introduces a large amount of variability to vowel realizations, making comparisons more difficult. It is thus concluded that further research on vowels in Afrikaans does well to primarily use wordlists instead of readings. Considering sociophonetic differences, it is concluded that there are clear differences in realizations between coloured and white speakers, and these differences are larger than any previously observed difference between geographically distinct groups of coloured speakers. The largest differences in vowel realization between white and coloured speakers pertains to the phoneme a, which is pronounced more closed and backed by the white speakers. Other notable differences between the groups were found for all diphthongs and for u, which is suggested to be fronted by the white speakers. Even though the results suggest that such differences exist, the small number of white and only female speakers in the data poses a problem regarding the generalizability of the results to the entire population of Afrikaans speakers. It is thus suggested that further research with a larger and more representative sample of white speakers is conducted.

In this study, cliff's index was introduced as a measure of distance between vowel spaces and compared with the overlap index. From this, it can be concluded that the indexes are correlated and coherently manages to capture differences between distributions. The measures are however differently influenced by certain aspects of differences. Cliff's index captures the distinctness of the vowel realizations in the vowel spaces to a larger extent than the overlap index, while the overlap index better captures differences in the variability of how the vowels are realized. Since Cliff's index is calculated based on Cliff's delta, the directionality of the distances can be found, making it in certain aspects a more useful measure of distance between single vowels in two vowel spaces, as it does not require a visual representation of the vowel spaces to interpret the results. This coupled with an insensitivity to outliers and strong correlation with shifts in central tendency suggests it to be a better primary measure for comparing vowel spaces than overlap index, since shifts in place of articulation should be a main concern of sociophonetic research. In conclusion, however, overlap index should if possible also be reported together with Cliff's index, as it could suggest differences in variability and modality which aren't necessarily captured by Cliff's index.

References

Baird, L., Evans, N., & Greenhill, S. J. (2022). Blowing in the wind: Using 'north wind and the sun' texts to sample phoneme inventories. *Journal of the International Phonetic Association*, *52*(3), 453–494. doi: 10.1017/S002510032000033X

Barber B., D. D., & Huhdanpaa, H. (1977). The quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.*, *22*(4), 398-403. doi: 10.1145/355759.355766

Britannica, E. (2023). *Cape province. Encyclopaedia Britannica.* Retrieved from `https://www.britannica.com/topic/cape` (December 16)

Britannica, E. (2024a). *Apartheid. Encyclopaedia Britannica.* Retrieved from `https://www.britannica.com/topic/apartheid` (December 16)

Britannica, E. (2024b). *Coloured. Encyclopaedia Britannica.* Retrieved from `https://www.britannica.com/topic/apartheid` (December 16)

Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin*, *114*(3), 494-509.

Eddy, W. (1977). A new convex hull algorithm for planar sets. *ACM Trans. Math. Softw.*, *3*(4), 398-403. doi: 10.1145/355759.355766

Jaccard, P. (1912). The distribution of the flora in the alpine zone.1. *New Phytologist*, *11*(2), 37-50. doi: 10.1111/j.1469-8137.1912.tb05611.x

Lobanov, B. M. (1971). Classification of russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, *49*, 606-608.

Meissel, Y. E., K. (2024). Using cliff's delta as a non-parametric effect size measure: An accessible web app and r tutorial. *Practical Assessment, Research, Evaluation,*, *29*(2). doi: 10.7275/pare.1977

Pastore, M., & Calcagnì, A. (2019). Measuring distribution similarities between samples: a distribution-free overlapping index. *Frontiers in Psychology*, *10:1089*.

Pfister Roland, J. M. D. R. F. J., Schwarz Katharina A. (2013). Good things peak in pairs: a note on the bimodality coefficient. *New Phytologist*, *4*. doi: 10.3389/fpsyg.2013.00700

R Core Team. (2023). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Wissing, D. (2020). Afrikaans. *Journal of the International Phonetic Association.*, *52*(3), 127-140. doi: 10.1017/S0025100318000269

Wissing, D., Baumann, A., & Pienaar, W. (2023). Die vokaalruimtes van bruinafrikaanse variëteite. *LitNet Akademies: 'n Joernaal vir die Geesteswetenskappe, Natuurwetenskappe, Regte en Godsdienswetenskappe*, *20*(3), 127-143.

5 Appendix

vowel_trans	vowel	example	type
ee	e	bees	gliding_monophthong
eu	ø	neus	gliding_monophthong
oo	o	boot	gliding_monophthong
ei	əi	byl	pure_diphthong
ui	œy	uit	pure_diphthong
œy	œy	uit	pure_diphthong
ou	œu	oud	pure_diphthong
ie	i	kies	pure_monophthong
uu	y	nuus	pure_monophthong
oe	u	koek	pure_monophthong
@	ə	kis	pure_monophthong
E	ɛ	ses	pure_monophthong
œ	œ	bus	pure_monophthong
O	ɔ	bos	pure_monophthong
o	ɔ	bos	pure_monophthong
ek	æ	ek	pure_monophthong
aa	a	baas	pure_monophthong
a	ɑ	bas	pure_monophthong

Figure 14: IPA table used to process files

Op 'n keer het Noordewind en Son stry gekry oor wie van hulle nou eintlik die sterkste was. Net toe kom daar 'n reisiger verby, gehul in 'n lekker warm mantel. Hulle besluit toe dat dié een wat dit kan regkry om die reisiger te dwing om sy mantel af te haal, die sterkste is. Toe blaas die wind ure lank vir al wat hy werd is. Maar hoe harder hy blaas, hoe digter hul die reisiger die mantel om hom. En ten laaste laat Noordewind al sy reuse pogings vaar. Toe skyn Son lekker warm, en dadelik trek die reisiger sy mantel uit. En so moes Noordewind erken dat Son die sterkste van hulle twee was.

Figure 15: The North Wind and the Sun in Afrikaans

Vowel	n_recordings	n_speakers	Mean_F1	Mean_F2
a	2075	300	513.6752	1769.9214
e	891	300	461.3580	1748.8967
i	566	300	552.2473	1796.5954
o	594	87	383.4074	899.0690
u	604	87	367.2003	953.8179
y	301	300	528.0166	1565.3754
ø	302	300	484.9205	1617.7285
œu	306	300	549.5556	1650.6405
œy	896	300	614.2288	1651.9196
ɑ	4170	300	530.0017	1741.4902
ɔ	764	87	438.0890	879.7461
ə	13889	300	560.9856	1722.9857
əi	4718	300	527.9290	1694.0659
ɛ	4173	300	503.9504	1852.7390

Figure 16: Dataset A aggregated over vowels

Vowel	n_recordings	n_speakers	Mean_F1	Mean_F2
a	1040	82	574.9885	1099.8144
e	419	82	456.7924	1450.1885
i	268	80	404.5224	1584.4291
o	477	82	402.7107	961.4361
u	545	82	386.6128	1043.3872
y	130	76	434.0154	1740.4000
ø	146	73	372.1301	1618.4521
œ	415	82	534.2145	1288.4554
œu	78	78	572.0897	1587.3718
œy	428	82	583.6776	1373.8598
ɑ	2081	82	583.2167	1213.6261
ɔ	669	82	443.3871	916.8580
ə	6193	82	487.8151	1328.0531
əi	2513	82	569.0501	1379.0243
ɛ	2248	82	547.4275	1430.8479

Figure 17: Dataset B aggregated over vowels

Vowel	n_recordings	n_speakers	Mean_F1	Mean_F2
a	229	215	768.3144	1440.1048
e	232	220	338.8233	2318.9957
i	233	220	345.0558	2271.5279
o	232	220	373.1078	755.8621
u	230	219	347.2304	809.5696
y	218	211	354.2294	2231.0505
ø	226	208	373.0619	2225.5000
œ	231	219	535.0693	1577.9654
œu	232	220	581.5819	1201.6379
œy	230	218	563.8087	1888.4043
ɑ	240	209	757.4625	1493.3083
ɔ	230	217	465.9609	838.2087
ə	228	216	524.3070	1665.9254
əi	227	215	595.4097	1769.2291
ɛ	231	219	461.4286	2096.7576

Figure 18: Dataset C aggregated over vowels

Vowel	n_recordings	n_speakers	Mean_F1	Mean_F2
a	18	9	609.9444	1105.4444
e	16	8	300.6250	2584.2500
i	18	9	305.9444	2511.4444
o	18	9	319.6667	797.7778
u	16	9	312.5000	1043.8125
y	18	9	273.5000	2454.2778
ø	18	9	298.8333	2591.9444
œ	18	9	413.2222	1812.9444
œu	18	9	650.1667	1395.8889
œy	17	9	533.0000	1603.5294
ɑ	18	9	756.4444	1339.8889
ɔ	18	9	376.2778	759.8333
ə	18	9	450.3333	2042.5556
əi	18	9	619.3889	1668.8333
ɛ	18	9	388.6111	2541.6111

Figure 19: Dataset D aggregated over vowels