**Sociophonetic variation in Afrikaans vowel production [DH/DS]**

In this project, students will analyze a dataset consisting of acoustic features of vowels (first and second formant) that have been produced by white and colored speakers of Afrikaans, a Germanic language that is spoken (mainly) in South Africa. The data have been recorded using two different methods (word lists vs. speech in context) by Daan Wissing (North West University, Potchefstroom, South Africa). Acoustic features have been already extracted for all configurations. The goal of the project is to compare the different configurations to assess (a) whether the examined sociolinguistic variants of Afrikaans differ from each other and (b) to what extent results based on different methodologies match.

**Some context:**

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), 127-146.

<https://www.litnet.co.za/die-vokaalruimtes-van-bruinafrikaanse-varieteite/>

**Data:**

* A: coloured Afrikaans, reading
* B: white Afrikaans, reading
* C: coloured Afrikaans, word list
* D: white Afrikaans, word list

**Task:**

1. Main questions:
   1. Sociolinguistic differences: A vs. B & C vs. D
   2. Methodological differences: A vs. C & C vs. D
2. Visualize each of the four vowel spaces A-D (F1 & F2); do this more beautifully than in the paper above!
3. Think of a neat way of comparing two vowel spaces (something along the lines of Wissing et al. 2023, or, even better, something different)
   1. How can one measure the distance (effect size) between two vowel spaces?
   2. How can one quantify to what extent vowel realizations are more dispersed on one of the two vowel spaces?
   3. Howe can one measure which vowel contributes most to the distance between two vowel spaces, i.e., which vowel is most impactful?

Beste Andreas

Dit was sinvol om die stelle terug te hou en weer deur te gaan vir foute - wat ek wel gekry het.

Hallo Andreas, hiermee die "regte" data, soos jy sien, is dit vier stelle, twee van bruin sprekers, en twee van wittes.

Die primêre vergelykings met betrekking tot ons navorsingsvraag is dié van A met B (beide Bruin sprekers) , en sekondêr C met D (beide Wit sprekers),   die datastelle van A en B is ewe groot, dit is ook deur dieselfde lesers gedoen. C en D is nie van dieselfde sprekers nie; D is ook groot, maar C is baie kleiner, wat vergelyking seker moeilik maak. Maar jy sal self sien as jy dit oopmaak.

Jy sal sekerlik self besluit om verskillende vergelykings, met bepaalde doelstellings, met mekaar te vergelyk (soos bv. B met D, wat dalk iets sal sê, as daar  'n kleiner verskil hier is as in die geval van B vergelyk met A). Natuurlik is ek baie nuuskierig, maar kyk maar waar jy die werk hieraan in jou program kan inpas.

Ek het geprobeer om sover moontlik eenvormigheid ten opsigte van die struktuur te handhaaf - soos wat ek met jou uitgeklaar het. Verder is al die vokale ook informeel benoem, dus nie met die amptelike IPA fonetiese simboolstelsel nie. Dis omdat sommige van die simbole nie aanvaarbaar is in  Statistica nie - ek eers daar gewerk en dit toe oorgeplaas Excel toe, wat problematies was.

Enkele opmerkings

      1. Ek het vir eers nie F1 gly en F2 gly ingesluit nie - dis soos wat ons ook met die Vokaalruimtes-artikel gedoen het, onthou?

      2. Die Vokale is nie in IPA-konvensie nie; dit kan later gebeur, voorlopig het ek dit so gedoen omdat Statistica en Excel nie lekker met mekaar daaroor wou gesels nie.

Sterkte hiermee; ek is baie nuuskierig wat Zaubermann hiermee gaan uitvoer!

Gegroet, eers weer.

D.

Hallo Andreas, hiermee die "regte" data, soos jy sien, is dit vier stelle, twee van bruin sprekers, en twee van wittes.

Die primêre vergelykings met betrekking tot ons navorsingsvraag is A met B (Bruin sprekers) , en sekondêr C met D (Wit sprekers),   A en B is van dieselfde groep bruin sprekers - dus ook ewe groot, maar C en D is twee stelle; D is ook groot, maar C baie klein. Maar jy sal self sien as jy dit oopmaak.

Jy sal sekerlik self besluit om verskillende vergelykings, met bepaalde doelstellings, met mekaar te vergelyk (soos bv. B met D, wat dalk iets sal sê, as daar  'n kleiner verskil hier is as in die geval van B vergelyk met A). Natuurlik is ek baie nuuskierig, maar kyk maar waar jy die werk hieraan in jou program kan inpas.

Ek het geprobeer om sover moontlik eenvormigheid ten opsigte van die struktuur te handhaaf - soos wat ek met jou uitgeklaar het. Verder is al die vokale ook informeel benoem, dus nie met die amptelike IPA fonetiese simboolstelsel nie. Dis omdat sommige van die simbole nie aanvaarbaar is in  Statistica nie - ek eers daar gewerk en dit toe oorgeplaas Excel toe, wat problematies was.

Gegroet, eers weer.

D.

Hallo Andreas, ek hoop en vertrou dat julle veilig terug is in W, en dat jy ook al weer jou voete kon vind. Dit was besonder lekker om julle weer te gesien het: Ur-lekker!

Soos belowe, het ek die omvattende datastel klaargemaak. Voordat ek dit vir jou stuur, wil ek net seker maak van die struktuur. Kyk asseblief die meegaande Excel-lêer. Jy sal sien:

      Die hoofvraag met betrekking tot die vergelyking van a) Woordelyste vs b) Lopende Teks (Die Noordewind en die Son) word gedek       as mens twee datastelle het, soos genoem in die eerste twee rye. Daar is 300 bruin sprekers wat sowel a) as b) gelees het. Dit is                  dit was ek gedink het om ons vraag aan te spreek. Die res is ekstra, en kan dalk interessant wees om ook te bekyk: ek het hier ook                nog b)-data van 80 sprekers van ''n ander groep, wat ons moontlik ook kan gebruik.

      Ek voeg WIT by, slegs interessantheidshalwe. Op die oomblik is ek net seker daarvan dat ek Woordelyste van 'n kleinerige                               groepe spekers het, maar moontlik ontdek ek ook nog Lopende Teks van dieselfde groep, of dalk van 'n ander groep. Maar dit lyk                 my nie juis bruikbaar nie. Moontlik is 'n sekondêre vergelyking van WIT se woordelys met dié van BRUIN tog wel handig. Wat sê jy?

My hoofvraag op die oomblik is of jy sal kan werk met één groot datastel met hierdie struktuur - dis soos ek jou verstaan het. As jy die groenlig, stuur ek alles vir jou.

Tot later, dan weer.

Baie groete daar

Daan.