The Control of Token-to-Token Variability: an Experimental and Modeling Study

Christine Mooshammer[1], Pascal Perrier[2], Susanne Fuchs[3], Christian Geng[1] and Yohan Payan[4]

[1]ZAS – Berlin, Germany, [2]ICP Grenoble, France, [3]QMUC, Edinburgh, UK, [4]TIMC, La Tronche, France

INTRODUCTION

Aim: Exploring factors influencing token-to-token variability during vowel production with:

- An experimental study based on EMMA, EPG, and acoustical data
- A modeling study using a 2D biomechanical tongue model

Investigated factors

Physical factors:

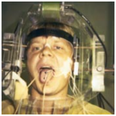
- Number of contacts between tongue and palate (EPG data)
- Intrinsic biomechanical properties of the tongue (Modeling)

Neurophysiological factor

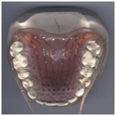
- Neural noise added to the motor commands (Modeling)

DATA ACQUISITION AND PROCESSING

Articulatory data consisted of simultaneous EMMA and EPG recordings.



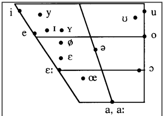
EMMA (AG100, Carstens)



EPG (Reading EPG3)

Material:

- CVCə nonsense words
- C: velar (e.g. gaake, gucke) or bilabial (e.g. baape, buppe)
- V: 14 German vowels



The German Vowel System

- Carrier sentence "Sage CVCə bitte" ("Say CVCə please")
- 10 repetitions
- Two conditions: with and without bite block (5mm thick between the second molars)
- Three speakers of German

Sensor positions and formant values were determined at the time where most of the sensor trajectories reached a turning point during the voiced portion of the vowel, and the fewest EPG contacts were observed (see Figure 1).

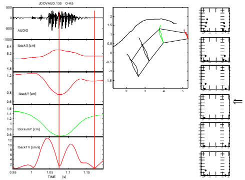
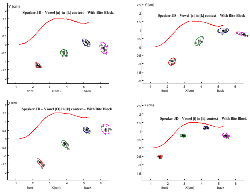


Figure 1: Measurement time on EMMA (left panel) and EPG (right panel) data

MEASURED ARTICULATORY VARIABILITY



General trends for the majority of cases:

For low vowels (see left panels on Figure 2) the main axes of the ellipses are essentially oriented in the high-front/low-back direction.

For high vowels (see right panels on Figure 2) the main axes of the ellipses are essentially parallel to the tongue contour at the sensor location.

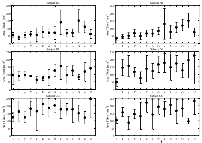


Figure 3: Averaged dispersion ellipses area (in mm^2) for the two rear transducers and for all subjects. Average values were computed from the data obtained in bite block and in normal conditions and in both consonant contexts.

Figure 3 shows that:

In five cases among six, vowel [i] has the smallest variability

The high front vowels [I, e, y, Y] show in the majority of cases less variability than low-back vowels like [a] or [O].

For two speakers vowel [u] has one of the largest variabilities among all vowels.

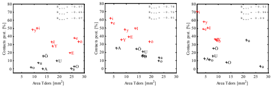
PALATAL CONTACTS AND ARTICULATORY VARIABILITY

The correlations between the number of palatal contacts and the sensor position (X and Y in the midsagittal plane) show for all speakers, that the higher and the more anterior the tongue position, the greater the number of palatal contacts.

In addition Figure 4 shows that :

For high vowels and for all speakers the amount of articulatory variability is clearly negatively correlated with the number of palatal contacts.

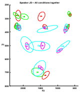
For low vowels there is no correlation for two speakers between the number of palatal contacts and the amount of articulatory variability.



Subject CG Subject DF Subject JD

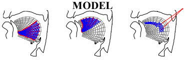
Figure 4: Relation between the Posteriority Index and ellipsis area for the sensor TDORS. Correlation coefficients are given for all vowels and for front and back vowels separately. Front vowels are depicted in red.

ACOUSTIC VARIABILITY

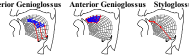


Acoustic variability tends to be larger for low vowels than for high ones = similar to articulatory data
BUT: Vowel [u]: small acoustic variability AND large articulatory variability ⇒ This non-linear relationship could influence speech control strategies in order to facilitate perception.

SIMULATIONS WITH A BIOMECHANICAL TONGUE MODEL



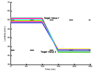
Posterior Genioglossus Anterior Genioglossus Styloglossus



Hyoglossus Verticalis Inferior Longitudinalis

Tested hypothesis: Articulatory variability at the targets is partly due to variability of the motor commands caused by the addition of a random error.

- The random error is modelled by the addition of a signal-related gaussian value to all motor commands (muscle threshold lengths, lambda) that specify the different targets of the speech sequence (See below).
- Vowels: [i], [a], [u]
- 20 simulations for each vowel



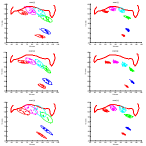


Figure 5: Generated dispersion ellipses simulated in velar consonantal context

Large noise (SNR=10dB): (1) Similar main ellipsis orientation for the three vowels along the high-front/low-back direction and (2) large variability with a minimum for [a] and a maximum for [u]

Smaller noise (SNR = 18dB): (1) Reduced variability with a minimum for [a] and a maximum for [u] and (2) main orientation essentially along the high-front/low-back direction with a tendency to be more horizontal for [a] than for the other two vowels.

The random alteration of the motor commands at the targets combined with tongue muscle insertions and anatomy could explain :

- the patterns of articulatory variability observed for low vowels like [a]
- the largest variability often observed for [u]

But can not explain

- the patterns of variability observed for high vowels.

CONCLUSION

Hypotheses explaining the observed patterns of token-to-token variability in vowel production:

(1) The variability originates in random errors on target motor commands.

(2) Patterns of variability are shaped by a combination of influences associated with neurophysiological and biomechanical aspects including palatal contacts (not yet implemented in the model); in the case of high front vowels, palatal contacts limit the variability in the vertical direction.

(3) For high back vowels like [u], additional studies are necessary, but the tendency for these vowels to be more variable could originate in the biomechanics of the tongue.