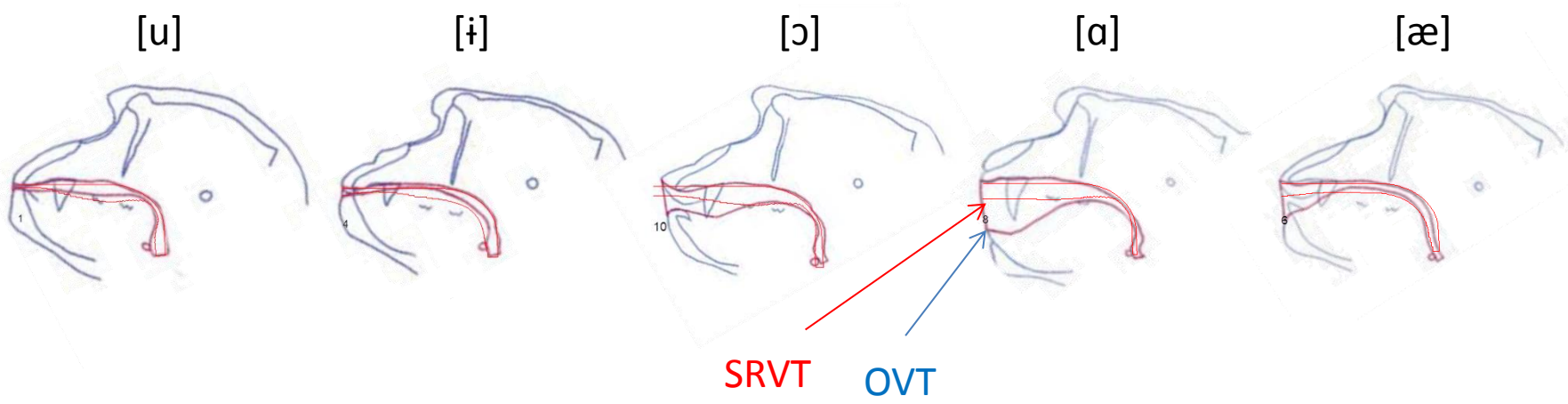
MONKEY VOCAL TRACTS ARE NOT SO “SPEECH READY“

Frédéric Berthommier

Univ. Grenoble Alpes, Grenoble INP, GIPSA-Lab, 38000 Grenoble, France



SCIENCE ADVANCES | RESEARCH ARTICLE

HUMAN EVOLUTION

Monkey vocal tracts are speech-ready

W. Tecumseh Fitch,[1,2]* Bart de Boer,[3] Neil Mathur,[4,5] Asif A. Ghazanfar[4,5,6]*

PRIOR

"Our findings imply that the evolution of human speech capabilities required neural changes rather than modifications of vocal anatomy. Macaques have a speech-ready vocal tract but lack a speech-ready brain to control it."

- Strong conclusions are formulated about the capacity of monkeys to produce vowels, thus to articulate phonemes with their upper vocal tract - in the context of an old and pationnate debate. –
- "Speech-ready" = Lack of brain control but the anatomy is set.

TOOL

A new generative model is used working at the area function level in the format proposed by Story and Titze (J. of Phonetics, 1998) (44 A + 1 dL). This can be seeded by real/human data or with a synthetic basis of area functions when these are not available as for the high-larynx VT of monkeys.

GOAL

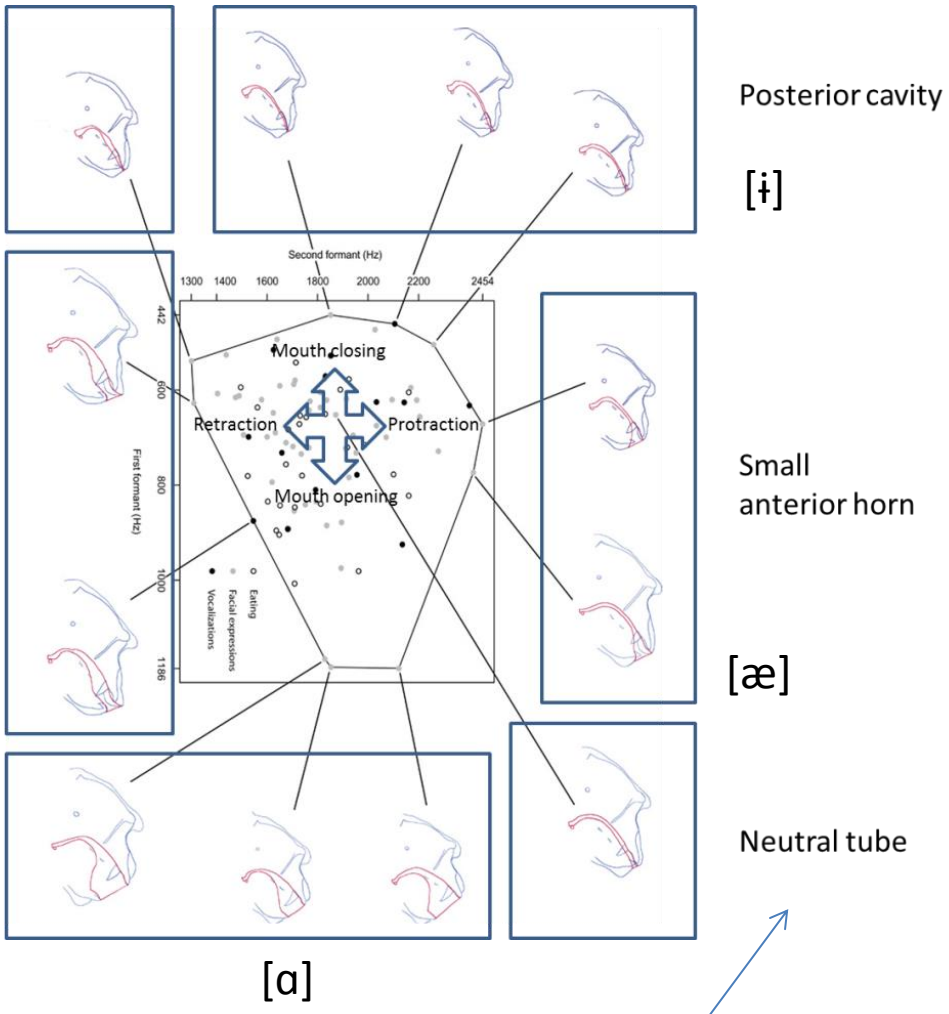
Falsification of this theory by setting of a "Speech-ready" and high-larynx monkey vocal tract (SRVT) and looking at the discrepancies with observations (OVT).

Setting of OVT

1) Repertoire of VT-Shapes associated with 2 gestural dimensions

Small set of published X-Ray sagittal cuts = Observed VT (OVT) (Fitch et al., 2016)

5 phonetic qualities (Boë et al., 2017)



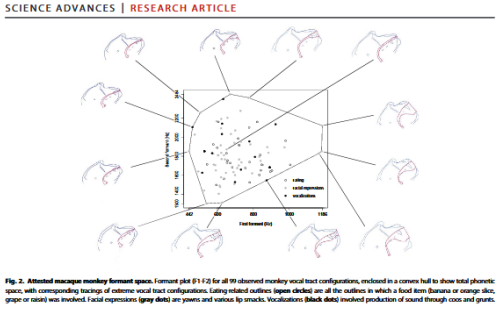
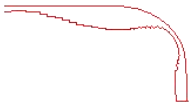


Fig. 2. Attested macaque monkey formant space. Formant plot (F1-F2) for all 99 observed monkey vocal tract configurations, enclosed in a convex hull to show total phonetic space, with corresponding tracings of extreme vocal tract configurations. Eating related outlines (**open circles**) are all the outlines in which a food item (banana or orange slice, grape or raisin) was involved. Facial expressions (**gray dots**) are yawns and various lip smacks. Vocalizations (**black dots**) involved production of sound through coos and grunts.

Features of the VT-Shape ("articulatory features") allowing F1-F2 formant resonances

Setting of SRVT

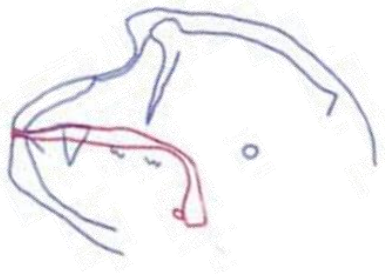
12th International Conference on Voice Physiology and Biomechanics
December 2-4, 2020, Grenoble, France

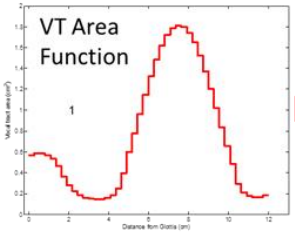


"Speech-Ready" Macaque (SRVT)

Reshaping

Articulatory modelling (VT-Shape)





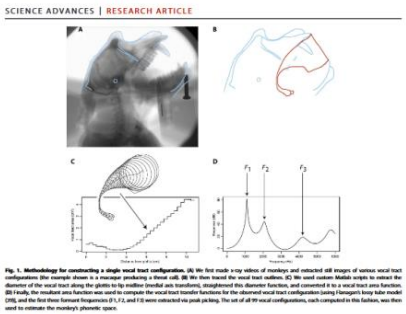
Acoustic modelling (Area function)

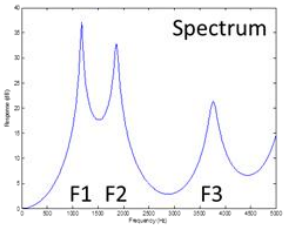
"Real" Macaque (OVT)

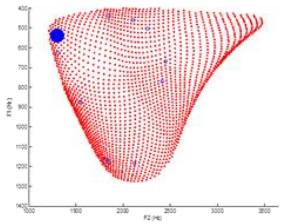
Automatic generation (Mathematical model)

Spectral output

Formants/Phonetic





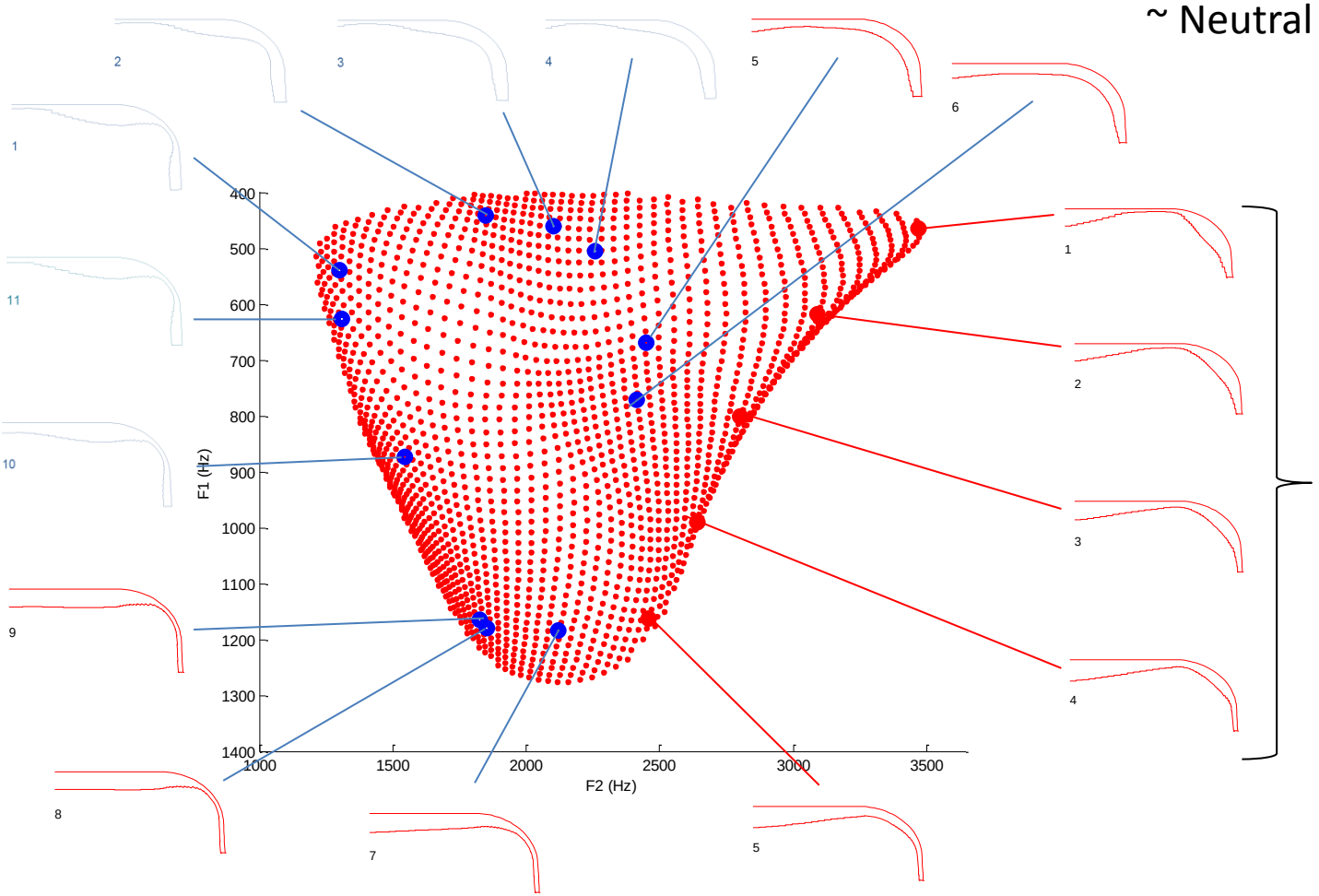


Output repertoire of SRVT

12th International Conference on Voice Physiology and Biomechanics
December 2-4, 2020, Grenoble, France

Region of OVT-SRVT similarities

~ Neutral of SRVT



Short and wide back cavity

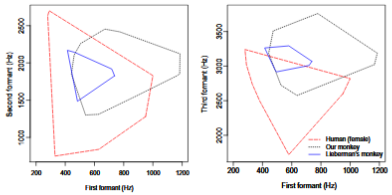


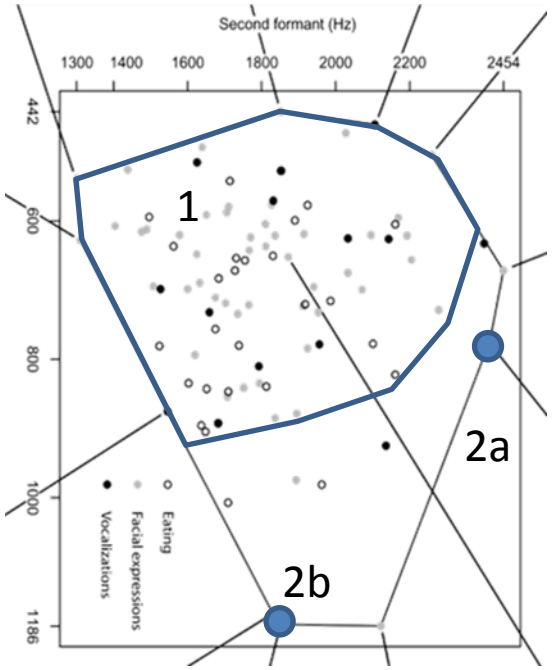
Fig. 3. Formant plot comparisons. Macaque monkey (**black dotted line**) versus human female vowel space (**red dashed line**; American English from Peterson and Barney (38)). The left panel shows F1 against F2, and the right panel shows F1 against F3. For comparison, the blue outline shows the previous macaque monkey estimates from Lieberman *et al.* (6).

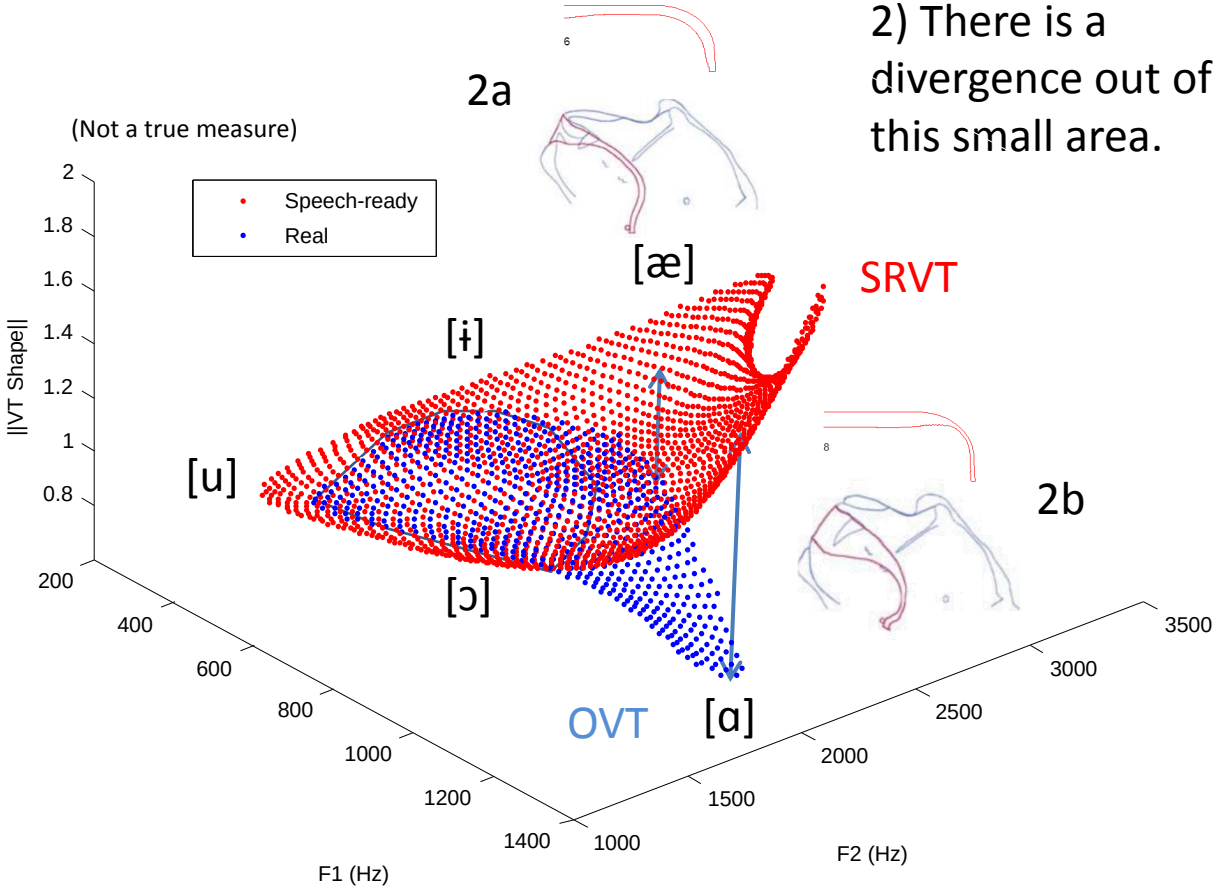
Lookup table available in SRVT.mat (to be read in Matlab with readSRVT.p)

OVT-SRVT discrepancies (I)

12th International Conference on Voice Physiology and Biomechanics
December 2-4, 2020, Grenoble, France

1) A small common area encloses a large majority of observed configurations.

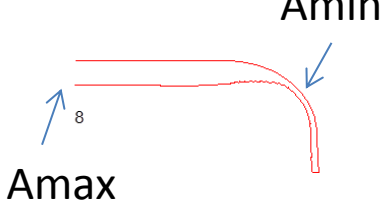




2) There is a divergence out of this small area.

2a) SRVT is in neutral configuration whereas OVT shows a small anterior horn producing a F2 increase.

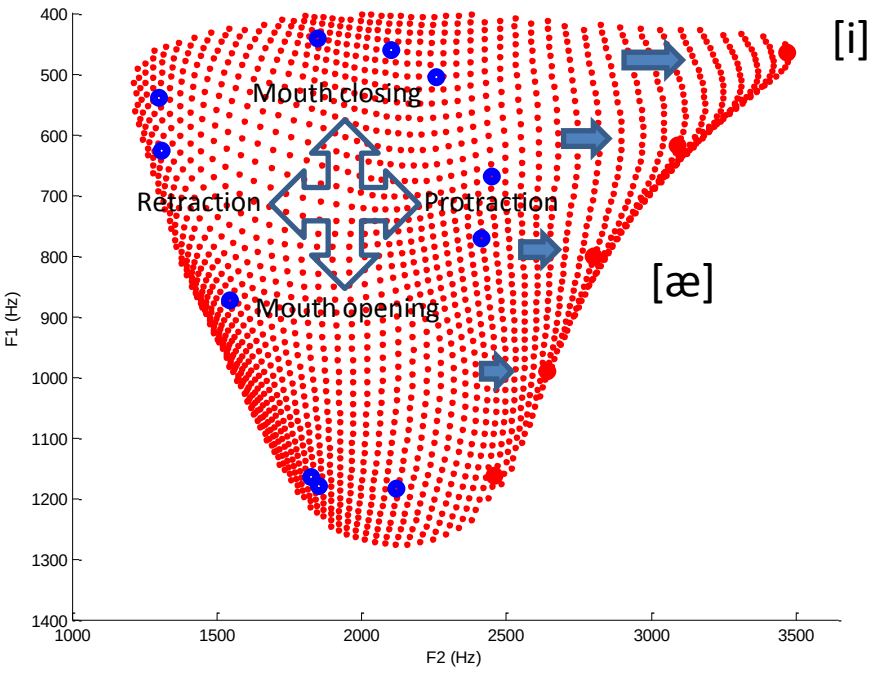
2b) The manner for increasing F1 via Amax/Amin is not the same: OVT is changing Amax only whereas SRVT is also modulating Amin.

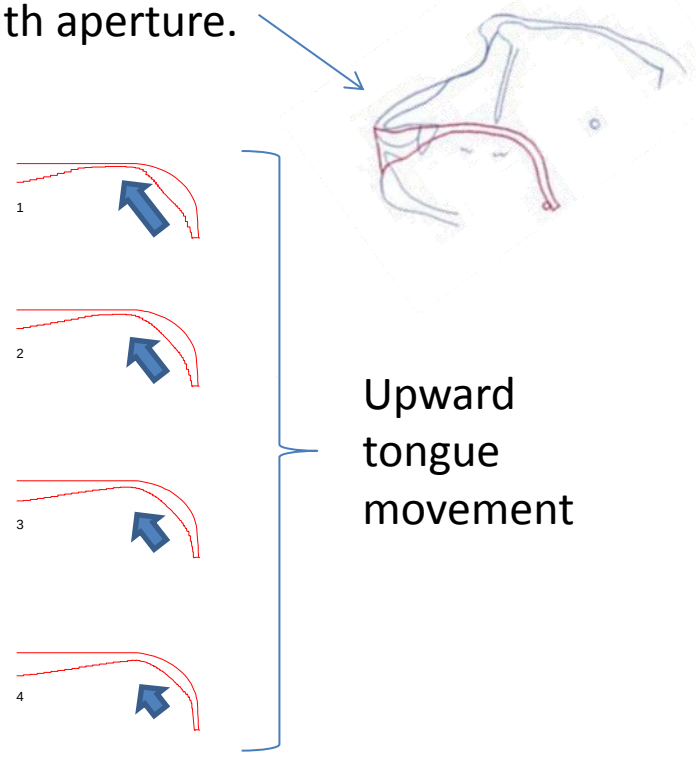


OVT-SRVT discrepancies (II)

3) A wide region of the F1-F2 plane is not reached by OVT.

SRVT maintains the lip opening whereas the tongue makes an upward movement. This forms a wide Helmholtz cavity together with an anterior constriction. This gesture depends on both larynx mobility and tongue configuration not present in monkeys. Presumably for anatomical reasons, these are only able to combine a mild/neutral protraction with an intermediate mouth aperture.





Link with the 5 phonetic qualities*

1) The OVT-SRVT overlap of configurations is small, mainly inside the [u] [ɨ] and [ɔ] region.

2) The 2 other phonetic qualities [ɑ] [æ] are not produced with the same manner of articulation:

2a - The [ɑ] phonetic quality is reached by OVT thanks to a wide mouth aperture whereas SRVT also makes a posterior constriction.

2b - The [æ] of OVT is determined by the shaping of the anterior part of the vocal tract. This is also consistent with the /wa/ segment of Boë et al. (2017). The animal is able to form a large anterior horn but the production of the SRVT [æ] also depends on a pronounced back articulation.

3) These back articulatory features are only present in the larger SRVT repertoire and they determine the production of the [i] phonetic quality.

*Described in: **Evidence of a Vocalic Proto-System in the Baboon (*Papio papio*) Suggests Pre-Hominin Speech Precursors**

Louis-Jean Boë, Frédéric Berthommier, Thierry Legou, Guillaume Captier, Caralyn Kemp, Thomas R. Sawallis, Yannick Becker, Arnaud Rey, Joël Fagot

Published: January 11, 2017 • https://doi.org/10.1371/journal.pone.0169321