UNIVERSITY OF CAMBRIDGE

CNBH, Physiology Department, Cambridge University
www.mrc-cbu.cam.ac.uk/cnbh

The relative contribution of glottal pulse rate and vocal tract length in size discrimination judgements. P59

Phil A. Gomersall, Thomas C. Walters, Richard E. Turner and Roy D. Patterson

pag36@cam.ac.uk **tcw24@cam.ac.uk** **ret26@cam.ac.uk** **rdp1@cam.ac.uk**

1. Introduction

•The perception of speaker size is related to both the glottal pulse rate of the speaker (GPR), which is perceived as pitch, and their vocal tract length (VTL).
• A speaker with a low GPR and long VTL (a man, for example) is perceived as larger than one with high GPR and short VTL (a child, for example).
• The percept of speaker size can therefore be visualised as a surface riding above the GPR-VTL plane which is high in the domain occupied by men and low in the domain occupied by children (Fig. 1).
•We wish to ascertain the configuration of this surface and thus find the trade-off between the two variables in making speaker-size judgements.
The size surface has already been mapped using absolute size judgements by subjects in a study by D. R. Smith (*This meeting, poster 60*).
However, **we present a novel, Bayesian method to generate an analytic form of the surface.**
A two-alternative-forced-choice experimental paradigm is used to infer the partial derivatives of the surface at specific points. These are then integrated to give the final size surface.

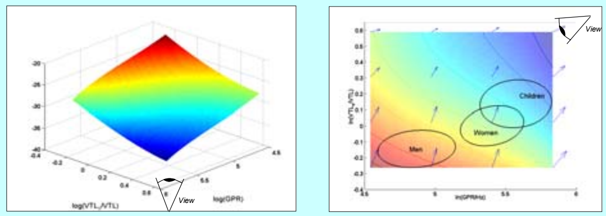

FIGURE 1 Right hand panel: A schematic of a candidate size surface (colour represents the perceived speaker size). Left hand panel: The gradients (blue arrows) and contours (coloured lines) of the candidate surface at 16 sample points. Ellipses (from the data of Peterson and Barney [1953]) show the positions of average men, women and children in the GPR-VTL plane.

2. Theoretical background

•The height of the perceptual size-surface (s) depends on the GPR (e^x) and VTL (e^y) of the speaker (Fig. 1)
•A perception of a speakers size ($\hat{s}$) is modelled as a noisy Gaussian sample taken from this surface:

$$P(\hat{s}(x,y)\mid s)=\frac{1}{\sqrt{2\pi\sigma_s^2}}\exp\left[-\tfrac{1}{2}\left(s(x,y)-\hat{s}(x,y)\right)^2\right]$$

•2-alternative forced-choice experiments (2AFC) necessarily monitor a listener's ability to perceive differences in stimuli, and can therefore be used to probe the partial derivatives of the size-surface.
•One trial of a 2AFC experiment consists of two intervals: In one a standard-sound is played and in the other a test sound.
•The sizes of the test and standard sounds can be chosen to be close together, such that the change in the height of the size surface between them is small and therefore well described by a 2D Taylor expansion:

$$s(x,y)=s(x_0,y_0)+\left.\frac{\partial s}{\partial x}\right|_{x_0,y_0}\Delta x+\left.\frac{\partial s}{\partial y}\right|_{x_0,y_0}\Delta y$$

•The variation in the JND for size across the GPR-VTL range is therefore represented by the local gradient of our size surface in the GPR and VTL directions
•Small JNDs mean larger partial derivatives and a steeper surface in that direction.

•3. Method

•The steepness of the size-surface at a point represents the JND and we want to select pairs of sounds on this surface which differ in size enough to be discriminable, and yet not as much as to break the criteria for the Taylor approximation.
•To infer the two partial derivatives of the surface at a point, we compare a standard stimlus to 8 test stimuli which are situated about it on the locus of an ellipse (Fig. 2).
•To encapsulate the variability of the JNDs between listeners and across the surface, the experiment was run on three sets of scaled ellipses: 1) Large for training, 2) Medium for first half. 3) Small for second half

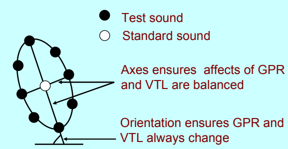

FIGURE 2: Positions of stimuli used to find the two partial derivates at the centre point.

•Each trial consisted of two stimuli, after which the subject had to choose which came from the smaller person.
•The stimuli consisted of four vowels selected at random from five canonical vowels spoken by a female which had been modified by the vocoder STRAIGHT to lie at the sample points.
•Seven subjects with normal hearing participated in the experiment, data is still being collected but so far 42 comparisons between each pair of vowel sounds have been made.

4. Results

•The results of the experiment are shown using a novel visual representation explained in Fig. 3
•A new Bayesian technique was developed to infer the partial derivatives and the errors on those inferences.
•The inferred partial derivatives of the size-surface are shown in Fig. 4.

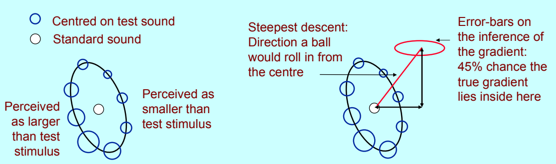

FIGURE 3 Example results: The area of the blue circles is proportional to the number of times the standard sound (positioned at the centre of the ellipse) was chosen as being smaller than this test sound.

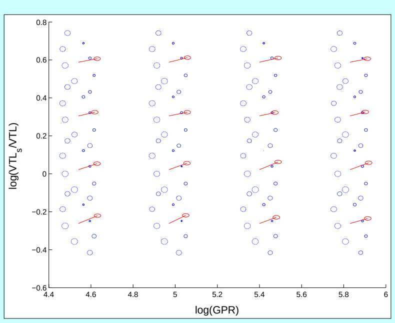

FIGURE:4: The experimental data (blue circles) and partial derivatives (red arrows) with error-bars (red ellipses). The relative magnitude of the vectors gives an indication of the steepness (the longer the vector, the steeper the gradient) of the size surface at the measured point

➤**In all cases both GPR and VTL affect the percept of size**

5. Analysis

To recover the size-surface the local gradient vectors should be interpolated and integrated. Interpolation of a function necessitates certain assumptions concerning the expected smoothness of the result to be fulfilled and general methods for interpolating are therefore hard to develop.
Bayes' Theorem, which requires these often tacit assumptions to be stated explicitly, incorporates such assumptions into a probabilistic framework as 'prior probabilities' and can therefore be used to develop a new method to recover the surface.
The assumptions used in fitting the surface are:
•The size-surface can be well described by a 2D polynomial over the sampled region.
•The size-surface will decrease with increasing GPR and SER
•There are no local extrema: all coefficients of the polynomial must be equal to, or less than zero.
The coefficients were found by maximising the Bayesian posterior probability (Fig. 5).
The upper and lower limits for the errors on the surface were deduced using a Monte Carlo (probabilistic sampling) technique. Laplace's approximation can also be used to put error-bars on the inferences.

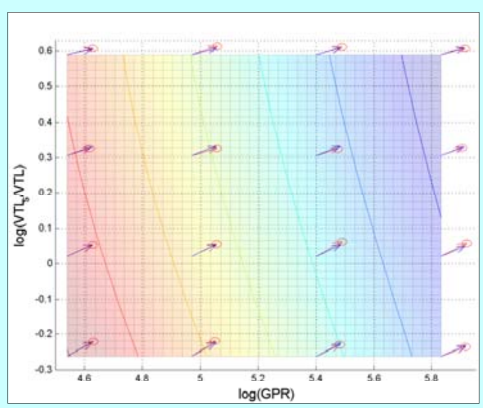

FIGURE 5: The inferred size-surface (colour), with contours of constant size (solid lines). The gradients inferred previously from the experiment are shown in red again with their error-bars (red ellipses). The analysis routine fits an n[th] order polynomial to the partial derivatives at the centre of each ellipse. In this case a 5[th] order polynomial was used. The gradients of this surface are shown at each sample point (blue arrows) for comparison.

6. Conclusion

➤**We have developed a novel psychoacoustic experiment and analysis to infer the size-surface.**
➤**The surface decreases with increasing GPR and decreasing VTL, and is roughly planar.**
➤**Small perturbations from higher order terms mean perception is not quite constant on the logarithmic of GPR and VTL.**

References

Peterson, G. E. and Barney, H. I., Control methods used in a study of vowels, *Journal of the Acoustical Society of America*, 24:75-184, 1952
Kawahara, H., Speech representation and transformation using adaptive interpolation of weighted spectrum: vocoder revisited. *ICASSP97*, volume 2:1303–1306, 1997
Mackay, D., *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 2003
Smith, D. R. R., The perception of sex and size in vowel sounds, BSA Poster 60, 2004

Acknowledgments

Research supported by the UK Medical Research Council (G9901257, G9900369)