ARCHIVES OF ACOUSTICS
Vol. **46**, No. 1, pp. 177–181 (2021)
DOI: 10.24425/aoa.2021.136571

Technical Note

Influence of Overpressure Breathing on Vowel Formant Frequencies

Milan VOJNOVIĆ(1), Miomir MIJIĆ(2)*, Dragana ŠUMARAC PAVLOVIĆ(2),
Nebojša VOJNOVIĆ(2)

(1) *Life Activities Advancement Center*
Belgrade, Serbia; e-mail: vojnovicmilan@yahoo.com

(2) *University of Belgrade*
School of Electrical Engineering
Belgrade, Serbia; e-mail: dsumarac@etf.rs, nebojsav83@hotmail.com
*Corresponding Author e-mail: emijic@etf.rs

(*received August 7, 2019; accepted October 5, 2020*)

Voice controlled management systems are based on speech recognition techniques. The use of such systems in combat aircraft is complex due to a number of critical factors which affect the accuracy of speech recognition, such as high level of ambient noise and vibration, use of oxygen masks, serious psycho-physical stress of speakers, etc. One of the specificity of the oxygen mask application is overpressure breathing. The results of the simulations presented in this paper show that the presence of overpressure on the order of 1000 Pa in the vocal tract has a significant influence on the first two formant frequencies. The formants discrimination field is significantly reduced when oxygen mask is used, influencing both perceptive and automatic discrimination of spoken vowels.

Keywords: mask; overpressure; transfer characteristic; vocal tract.

1. Introduction

Voice communication in the cockpit of a combat aircraft is characterized by a high level of ambient noise and by the use of the oxygen mask. The use of the mask affects speech with several aspects: changing the acoustic structure of the vocal tract, influencing articulation, changing the breathing process, changing audio-perceptive feedback, etc. Speech in the cockpit of the aircraft introduces one specific change that reflects on the vocal tract, i.e. the overpressure in the mask.

The overpressure causes a physical deformation of the vocal tract and the mask's chamber. The overpressure values can reach up to 60 mmHg (8000 Pa) in case of a great $+G_Z$ load (Rainford, Gradwell, 2006). During high-altitude flights, the overpressure value is around 30 mmHg. The influence of overpressure on the vocal tract is already described in (South, 2001). It has been shown that the first two formant frequencies tend toward values $F_1 \approx 550$ Hz and $F_2 \approx 1570$ Hz. These frequencies are approximately equal to the first two quarter-wave resonances of a uniform, cylindrical tube, 17 cm long and open at one end.

Upon closing of the vocal tract using an oxygen mask, as in the case of combat aircrafts, two important effects can be observed: the vocal tract's effective length is increased by the length of the mask, and the vocal tract is no longer a tube open at one end, but a tube closed at both ends. There are no data in the literature describing the consequences overpressure has on speech when using an oxygen mask. In this paper we present the results of the analysis regarding variations in formant frequencies which are caused after the closure of the vocal tract using an oxygen mask, and by the overpressure arising in the mask during breathing. The analysis is conducted through a simulation using an equivalent electrical circuit of the vocal tract (Fant, 1970; Flanagan, 1972).

2. Analysis description

Transfer function of the vocal tract is a function of its physical shape and can be determined using the analogy theory (Badin, Fant, 1984). For this purpose, an acoustical model of the vocal tract is utilized, comprising a series of connected, short cylinders of various diameters. Common segment lengths are 5 to 10 mm. Based on the analogy theory, short, uniform cylindrical tube is modeled using a symmetrical T-shaped four-port network. Elements of each four-port network are defined using the length and radius of the cylinder it represents (Flanagan, 1972). The final equivalent circuit of the vocal tract is obtained after cascading all the T-shaped circuits. Transfer function of thusly obtained equivalent electrical circuit is in fact the transfer function of the vocal tract where the maxima represent formant frequencies.

In the electrical model of the vocal tract, acoustical inductance and capacitances have a primary influence on the transfer function's resonant frequencies (vowel formant frequencies) (Fant, 1970; Flanagan, 1972). Acoustical resistances and conductances influence the resonance bandwidths. According to the relations defining the electrical model of a short cylindrical tube, two parameters are relevant for the calculation of the acoustical inductances and capacitances: speed of sound and air density.

Vocal tract walls are soft tissues of finite rigidity. An increase of the pressure inside of the vocal tract makes it "inflate", thus changing its shape and its cross-section. The same principle can be applied to the mask's chamber, since its walls are made of elastic materials. Based on the results of experimental measurements involving cylindrical tubes, it has been shown that there exists a linear relation between the tube pressure and the variation of its radius (Svirsky *et al.*, 1997). That relation can be described with the following equation:

$$\Delta r = pC, \tag{1}$$

where Δr denotes the variation of the tube radius, p denotes the pressure inside the tube, and C stands for the elasticity of the tube wall. An increase of pressure (p_{OVER}) inside a cylindrical tube with a radius r_0, causes an increase of the tube's cross-section, given by:

$$A_{\text{OVER}} = \pi(r_0 + \Delta r)^2 = \pi(r_0 + p_{\text{OVER}}C)^2. \tag{2}$$

Here, r_0 denotes the starting tube radius with the overpressure being $p_{\text{OVER}} = 0$. There are also experimental results of the vocal tract wall elasticity given in (Ernsting, 1966; Svirsky *et al.*, 1997). The average value of the elasticity can be approximated as $C = 2 \cdot 10^{-4}$ cm/Pa.

The analysis of the overpressure influence on the vocal tract formant frequencies is conducted using relation (2). Therefore, the variation of the vocal tract shape is estimated in the presence of overpressure during breathing. Vocal tract configurations were taken from the literature (Fant, 1970) for the case of Russian vowel pronunciation. A lossy vocal tract model was used (due to thermal and viscosity losses), and the radiation impedance of the mouth opening is modeled using an electrical circuit simulating the radiation of a piston mounted on a sphere (Vojnović, Mijić, 2005). The mask volume used in the simulations was 280 cm^3, shaped like a calotte with a maximal diameter of 10.7 cm and height 4 cm. This size and shape are approximately the same as in pilot masks used in military aircrafts. In the simulations, we assumed a non-radiating mask, i.e. with no input or output valves.

The overpressure inside the vocal tract was varied from 0 to 1000 Pa with a 200 Pa step. Vowel formant frequencies were calculated for six values of overpressure. The results of the analysis are related to the overpressure variations up until 1000 Pa, although these variations in real-life circumstances, in battle aircrafts, can be even higher.

3. Results of the analysis

Figure 1 shows the F$_1$–F$_2$ dependence of the analyzed vowels on the overpressure. Each symbol represents a single overpressure value which was varied from 0 to 1000 Pa, with a 200 Pa step. Thin dotted lines show the general tendency of the formant frequency variation upon increasing the overpressure. These lines intersect in the vicinity of the point defining the first formant frequency around 780 Hz, and again close to the second formant frequency around 1450 Hz (circle in Fig. 1). It is shown that after increasing the vocal tract overpressure, formant frequencies of all vowels tend toward a single, unique value, which means that the vocal tract shape tends toward a unique shape. Due to the existence of the overpressure inside the vocal tract, an "inflating" effect can be observed, and hence all vocal tract configurations obtain a similar shape no matter which vowel was pronounced.

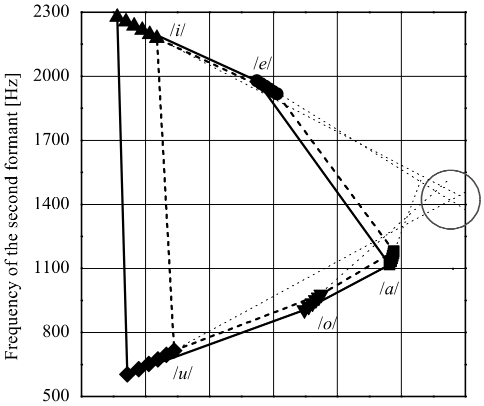

Fig. 1. F$_1$–F$_2$ diagram based on the overpressure variation inside the vocal tract.

For a tube 21.5 cm long and closed at both ends, the first and second resonant frequencies are around

790 Hz and 1580 Hz, respectively. In the available data on Russian vowels, the vocal tract length is 17–18 cm. If we suppose that a mask depth of 4 cm is used in the analysis, we have agreement with the demonstrated calculations. Presence of overpressure leads to the vocal tract obtaining the shape of a uniform cylindrical tube closed at both ends, and therefore formant frequencies tend toward the appropriate value.

According to Fig. 1 and the F_1–F_2 field, the range of the formant frequency variation is narrowed after increasing the overpressure inside the vocal tract. That range is highlighted with a dashed line, compared to the starting values which are denoted using solid lines. This narrowing effect can influence the perceived precision of the pronounced vowel, since the discriminatory areas around the vowels are smaller. It should be noted that the analysis includes only the pressure variation of 1000 Pa, and in real-life scenarios pressure variations can be greater. With this kind of overpressure, the formant frequency variation would be even more drastic, and the discriminatory vowel field would be significantly disturbed.

The influence of overpressure can be best appreciated by looking at the lower formant frequencies, namely the first and the second one. For the overpressure value of 1000 Pa, the frequencies of the first vocal formant increase up to 27% compared to the case with no overpressure. For the same variation in overpressure, second formant frequencies vary from −4% to +18%. For greater values of overpressure, formant frequency variations are even larger. It has been checked that for overpressure equal to 3000 Pa, first formant frequency is increased for almost 60%, and the range of second formant frequency variation is from −10% to 40%. It is interesting to note that variations in both first and second formant frequency tend toward the same values regardless of the spoken vowel. This is shown in Fig. 1 with dashed lines and a circle denoting the aforementioned target value.

It can be seen from Fig. 1 that the presence of overpressure causes greater variations in high vowels. In these vowels, the configuration of the vocal tract implies small cross-section area of either the mouth opening or the part of the vocal tract at the "tongue-palate" region. For the analyzed Russian vowels, the cross-section area of the mouth opening during pronunciation of the vowel /*u*/ equals 0.65 cm^2, which is the same as the cross-section area of the "tongue-palate" region during pronunciation of the vowel /*i*/. The variations in the cross-section area caused by overpressure, according to Eq. (2), are more pronounced at narrower parts of the vocal tract. Hence, configurations of the vocal tract containing such narrow areas, will generally suffer greater shape variation, which in turn implies larger formant frequency variation.

According to the obtained results, a general behavior of the formant frequency variation cannot be deduced with great certainty. With some vowels, first formant frequency (or second) decreases along with the increase of overpressure. With other vowels, the observed behavior is quite the opposite. The most evident fact is that the overpressure has the largest effect on the first two formant frequencies. The first formant frequency always increases along with the overpressure. The second formant frequency also increases along with the overpressure, except in the case of front vowels (/*e*/ and /*i*/).

Percentage variations of the formant frequencies of some vowels are extremely large. For example, first formant frequency of vowels /*u*/ and /*i*/ is increased by more than 20% with the 1000 Pa overpressure. The greatest formant frequency variations have been measured with the vowels /*u*/ and /*i*/, and somewhat lesser variations (below 10%) with the remaining three vowels. As a matter of fact, the variations caused by overpressure are nonlinear. Greater variations are mostly seen with lower formants. The greatest variations have been measured with the first formant. Variations of the second formant frequencies are less pronounced, while the variations regarding the remaining formants are negligible. If the overpressure is increased up to 3000 Pa, the percentual variations in formant frequencies are drastically larger. For example, with the vowels /*u*/ and /*i*/, the variations are on the order of 40%, and 10% with the remaining vowels.

In the presented analysis only the variations in the shape of the vocal tract are taken into account, opposed to the variations in the shape of the mask chamber. With a similar analysis it can be shown that the vowel formant frequencies will slightly decrease if only the mask shape is varied due to the presence of overpressure. Variations are mostly related to the first two formant frequencies and are less than 1%, referenced to the case with no overpressure. Basically, the formant frequencies approach those values corresponding to normal speech, i.e. speech without an oxygen mask. The presence of overpressure increases the volume of the mask chamber, and thus reduces the influence of the mask on the vowel formant frequencies. Overpressure, in a way, cancels the mask's influence on speech, although in a symbolical amount.

If we analyze simultaneous influence of the overpressure on the shapes of both the vocal tract and the oxygen mask, we obtain results differing negligibly from those given in Fig. 1.

4. Discussion

Presented simulation results generally agree with the results of the researches done by other authors (South, 2001). In these investigations, "centralization" of the formant frequencies caused by overpressure, is also determined. In (South, 2001), a model of the vocal tract without an oxygen mask is used, hence formant frequencies tend toward $F_1 \approx 550$ Hz and $F_2 \approx 1570$ Hz. These frequencies are approximately equal to the first two quarter-wavelength resonances of a uniform, cylindrical tube, open at one end, and 17 cm long (average vocal tract length in an adult male). As shown in this paper, when an oxygen mask is included in the vocal tract model, its "effective" length is increased by the mask length, and therefore it is no longer shaped like a tube open at one end, but as a tube closed at both ends.

Figure 2 depicts percentage frequency variations of the first three formants of the Russian vowel /*a*/ when the speed of sound is varied in the range from −5% to +5% compared to the nominal value ($c = 35\,300$ cm/s, at the temperature of 37°C). In the figure, the variations for the first three formant frequencies are plotted. The influence of the speed of sound is practically the same for all three formant frequencies, hence a single line is visible. The same functional dependency applies for the remaining vowels from Fig. 2. There is a linear relation between the percentage variation in vowel formant frequencies and the percentage variation in the speed of sound. If we change the speed of sound by a certain percentage, all the formant frequencies are thus changed by the same amount. The functional relation between the speed of sound and the formant frequencies is independent of the vocal tract shape, since same results (as those shown in Fig. 2) were obtained for all five Russian vowels.

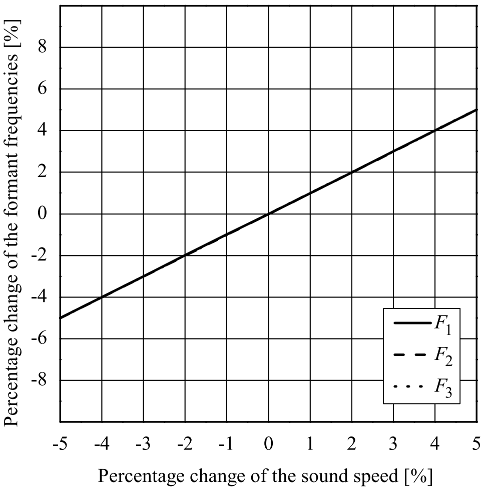

Fig. 2. First three formant frequency variations in the vowel /*a*/ as a function of the variation in the speed of sound (lines overlap).

If pure oxygen is used for breathing instead of air, the speed of sound is altered. At room temperature, speed of sound equals 343 m/s in air and 317 m/s in oxygen, which is a difference of 7.5%. From Fig. 2 it can be seen that the formant frequencies, in that case, decrease by the same amount. Thus, breathing pure oxygen, to some extent, compensates the influence of overpressure on the vowel formant.

Figure 3 shows the percentual changes of the first three formants of the Russian vowel /*a*/ when air density is varied in the range from −5% to +5% compared to the nominal value ($\rho = 1.14 \cdot 10^{-3}$ g/cm^3).

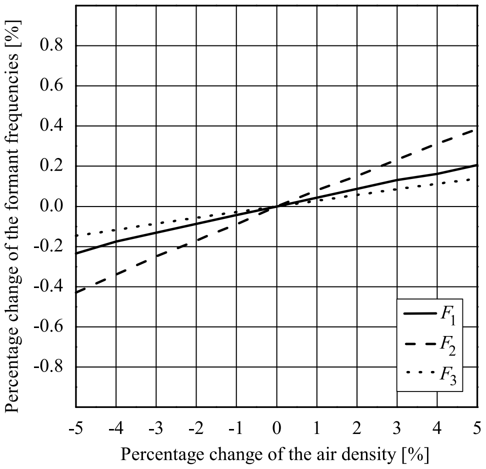

Fig. 3. Percentual variations of the formant frequencies in the Russian vowel /*a*/ as a function of the percentual variation of air density.

The variation tendency is the same in Fig. 3 as in Fig. 2: reducing air density decreases all the formant frequencies. The range of the formant frequency variation is around ten times smaller when air density is varied, compared to the case when speed of sound is varied. For example, if we reduce air density by 5%, compared to the nominal value, formant frequencies will decrease by around 0.5%. The variations in formant frequencies depend also on the vocal tract shape, since the diagram from Fig. 3 is not the same for all vowels.

Oxygen density is around 10% higher than air density, so (according to Fig. 3) vowel formant frequencies will increase by around 1%. If we look only at the increase of air density, due to the presence of overpressure, those variations are significantly lower and can be disregarded. Namely, for each 1000 Pa overpressure increase, air density increases for around 1%, which in turn causes an increase of the formant frequencies for around 0.1%.

5. Conclusion

In this paper we considered the influence which overpressure breathing has on vowel formant frequencies. Unlike all the investigations done so far by other authors, we analyzed all aspects of overpressure breathing using an oxygen mask: the effect of closing the vocal tract using an oxygen mask, the influence overpressure has on air density, and also on physical alterations of both the vocal tract and the mask, and finally the effect of using pure oxygen for breathing instead of air.

Presented simulation results show that breathing in the presence of overpressure has a significant impact on speech, i.e. on the formant structure of the pronounced phoneme. Variations in formant frequencies come as consequences of an altered vocal tract shape. Due to overpressure, vocal tract "inflates". The same effect can be observed in a mask, but the variations in the mask shape have no significant influence on the vowel formant frequencies.

Presence of overpressure, which causes the vocal tract to change its shape, has a dominant effect on vowel formant frequencies. Nevertheless, variations in formant frequencies caused by overpressure in the vocal tract, are significant and have to be considered when voice control systems are designed.

Variations in air density caused by overpressure are negligible and can be disregarded. If pure oxygen is used for breathing, then, due to higher density, vowel formant frequencies are increased for around 1%, but are simultaneously decreased for around 5% due to the decrease of speed of sound in oxygen (compared to air). Thus, upon inhaling oxygen, overpressure influence is only partly compensated.

Acknowledgements

This research was supported by grants TR32032 from the Ministry of Science and Technological Development of the Republic of Serbia.

References

1. Badin P., Fant G. (1984), Notes on vocal tract computation, *STL-QPSR*, **25**(2–3): 53–108, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm.
2. Ernsting J. (1966), Some effects of raised intrapulmonary pressure in man, *AGARD Monograph, AGARDOGRAPH 106*, Technivision Ltd., Maidenhead.
3. Fant G. (1970), *Acoustic Theory of Speech Production*, Mouton, The Hague.
4. C (1972), *Speech Analysis, Synthesis, and Perception*, Springer-Verlag, New York.
5. Morse P.M. (1986), *Vibration and Sound*, Ed. by Acoustical Society of America.
6. Morse P.M., Ingard K.U. (1968), *Theoretical Acoustics*, McGraw-Hill, New York.
7. Rainford D., Gradwell D. [Eds] (2006), *Ernsting's Aviation Medicine*, 4th ed., Hodder Arnold, London.
8. South A. (2001), A model of vowel production under positive pressure breathing, *Proceedings of EUROSPEECH-2001*, pp. 1515–1518, Aalborg, Denmark.
9. Stevens K.N., Kasowski S., Fant G. (1953), An electrical analog of the vocal tract, *The Journal of the Acoustical Society of America*, **25**(4): 734–742, doi: 10.1121/1.1907169.
10. Svirsky M., Stevens K., Matthies M., Manzella J., Perkell J., Wilhelms-Tricarico R. (1997), Tongue surface displacement during bilabial stops, *The Journal of the Acoustical Society of America*, **102**(1): 562–571, doi: 10.1121/1.419729.
11. Vojnović M., Mijić M. (2005), An improved model for the acoustic radiation impedance of the mouth based on an equivalent electrical network, *Applied Acoustics*, **66**(5): 481–499, doi: 10.1016/j.apacoust.2004.09.002.
12. Wakita H., Fant G. (1978), Toward a better vocal tract model, *STL-QPSR*, **19**(1): 929, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm.