

Uticaj natpritiska u vokalnom traktu na formantne frekvencije vokala

Milan Vojnović

Sadržaj — Sistemi za upravljanje govorom bazirani su na tehnikama prepoznavanja govora. Primena sistema za upravljanje govorom u borbenim vazduhoplovima je posebno problematična zbog niza nepovoljnih faktora koji utiču na tačnost prepoznavanja govora: visok nivo ambijentalne buke i vibracija, korišćenje kiseoničke maske, velika psiho-fizička naprezanja govornika i dr. Jedna od specifičnosti korišćenja kiseoničke maske je disanje sa natpritiskom. Rezultati simulacija pokazuju da prisustvo natpritiska u vokalnom traktu reda 1000 Pa ima značajan uticaj na frekvencije prva dva formanta vokala. Promene frekvencija ovih formanta su oko dvadesetak procenata u poređenju sa slučajem kada nema natpritiska.

Gljučne reči — formantne frekvencije vokala, kiseonička maska, prepoznavanje govora, upravljanje govorom.

I. UVOD

GOVOR je jedan od vidova razmene informacija između ljudi. On se može uspešno primeniti i kod komunikacije na relaciji čovek-mašina. Obično se ta komunikacija svodi na izdavanje različitih govornih komandi od strane čoveka i "obaveštavanje" od strane mašine o tome kako je izrečena komanda shvaćena. Sistemi za prepoznavanje govora (Speech To Text – STT – Govor U Tekst) omogućuju da se govorne komande čoveka pretoče u konkretne programske i/ili mehaničke akcije neophodne za upravljanje mašinom. Sa druge strane, sistemi za sintezu govora (Text To Speech – TTS – Tekst U Govor) omogućuju obaveštavanje čoveka o tome kako je mašina "razumela" zadatu komandu i to na način najprihvatljiviji za čoveka – pomoću govora. Sistemi obaveštavanja, pored povratne informacije o prijemu komande, nude i mnoge druge aspekte obaveštavanja koji omogućuju jasan i potpun uvid u stanje procesa koji se kontroliše kao i informacije o raspoloživim resursima mašine.

Upravljanje govorom je našlo veliku primenu u slučajevima kada je manipulativna sposobnost čoveka redukovana bez obzira da li je u pitanju hendikep, ili on obavlja takvu vrstu posla koja podrazumeva upravljanje istovremeno sa više sistema ili uređaja. U takvim situacijama najčešće se pribegava upravljanju pomoću govornih komandi. Tipičan primer gde se koristi upravljanje govorom je borbeni vazduhoplov. Letenje

borbenim vazduhoplovima podrazumeva da je desna ruka pilota na pilotskoj palici, a leva na komandama gasa. Kontrola vazduhoplovnih sistema i uređaja zahteva fizičku manipulaciju različitim prekidačima, selektorima, tastaturama i sl., odnosno podizanje ruke pilota sa pilotske palice ili ručice gasa. Ovo je posebno problematično tokom leta u složenim meteo uslovima, ili u fazi leta neposrednog borbenog dejstva. U tim situacijama pažnja pilota treba da bude usmerena na upravljanje vazduhoplovom, a svaka potreba za promenom režima rada različitih vazduhoplovnih sistema i uređaja remeti proceduru upravljanja vazduhoplovom. Pre tridesetak godina razvijen je sistem HOTAS za borbene vazduhoplove (Hands On Throttle And Stick - Ruke na komandama gasa i pilotskoj palici). Sistem HOTAS podrazumeva da su ruke pilota stalno na pilotskoj palici i ručici gasa, a da se upravljanje vazduhoplovnim sistemima i uređajima obavlja govornim komandama.

Osnovu sistema za upravljanje govorom čini uređaj za prepoznavanje govora koji vrši konverziju govornog signala u tekst. Za realizaciju ovih uređaja koriste se različite tehnike i algoritmi koji zavise od: zahtevane tačnosti prepoznavanja govora, fonda reči koje treba prepoznati, ambijentalnih uslova u kojima se obavlja govorna komunikacija i drugog. Uslovi koji vladaju u kabini borbenog vazduhoplova su krajnje nepovoljni sa stanovišta govornih komunikacija. Specifičnosti govornih komunikacija u borbenim vazduhoplovima su: visok nivo ambijentalne buke, visok nivo vibracija, korišćenje kiseoničke maske, velika psiho-fizička naprezanja govornika (pilota) i dr. Upotreba kiseoničke maske je posebno interesantna jer ona utiče na govor sa više aspekata: menja akustičku strukturu vokalnog trakta, menja artikulaciju, menja proces disanja, menja audio-perceptivnu povratnu spregu, nameće korišćenje elektro-akustičkih pretvarača (mikrofona i slušalica) ...

U situacijama povećanog fizičkog naprezanja pilota (na primer velika ubrzanja vazduhoplova), disanje se pospešuje tako što se koristi čist kiseonik umesto vazduha, a on se u pilotsku masku dovodi pod određenim pritiskom. Ovaj povećani pritisak (natpritisk) takođe može da bude uzrok promena govornog signala. Disanje sa natpritiskom dovodi do fizičke deformacije vokalnog trakta i same komore maske što ima posredan uticaj na govor, odnosno na formantnu strukturu izgovorene foneme. Standardna šema dovođenja natpritiska u funkciji +Gz ubrzanja je: natpritisk se povećava za 10 mmHg za svako povećanje +Gz ubrzanja od 1g i to počevši od 2g gde je natpritisk 4 mmHg. Prema ovoj šemi dobija se sledeće:

Milan Vojnović, Institut za eksperimentalnu fonetiku i patologiju govora, Gospodar Jovanova 35, Beograd, Srbija (e-mail: vojnovicmilan@yahoo.com).

4 mmHg (533 Pa)	za +Gz ubrzanje od 2g,
14 mmHg (1866 Pa)	za +Gz ubrzanje od 3g,
24 mmHg (3199 Pa)	za +Gz ubrzanje od 4g,
34 mmHg (4532 Pa)	za +Gz ubrzanje od 5g,
44 mmHg (5865 Pa)	za +Gz ubrzanje od 6g itd.

Kod realizacije sistema za prepoznavanje govora koji se koriste u borbenim vazduhoplovima moraju se uzeti sve napred navedene specifičnosti vazduhoplovnih govornih komunikacija. Delikatnost projektovanja ovih sistema upravo proističe iz kompleksnosti sredine u kojima oni treba da rade pri čemu se zahteva izuzetno visoka tačnost rada sistema (tačnost prepoznavanja reči preko 99%).

U ovom radu je analiziran samo jedan od mnogih vidova uticaja maske na govor: uticaj natpritiska na formantnu strukturu vokala.

II. POSTAVKA PROBLEMA

Disanje sa natpritiskom utiče na tri stvari relevantne za govor:

- menja se oblik vokalnog trakta i kiseoničke maske,
- menja se gustina vazduha i
- menja se brzina zvuka zbog korišćenja kiseonika umesto vazduha.

U ovom radu razmatran je samo prvi aspekt uticaja natpritiska na govor: kakve su promene formantne strukture vokala kada se menja oblik vokalnog trakta i maske zbog prisustva natpritiska.

Između fizičkog oblika vokalnog trakta i njegove prenosne karakteristike postoji jednoznačna veza. Ako se znaju gabaritne dimenzije vokalnog trakta uvek se može odrediti njegova prenosna karakteristika. Dobijanje prenosne karakteristike vokalnog trakta, na osnovu njegovog oblika, postiže se korišćenjem teorije analogija [1]. Na osnovu ove teorije, kratka uniformna cilindrična cev se u električnom domenu modeluje simetričnim T-četvoropolom. Elementi ovog T-četvoropola su definisani dužinom i poluprečnikom cilindrične cevi [2]. Da bi se vokalni trakt modelovao ekvivalentnim električnim kolom on se prvo aproksimira (segmentira) kratkim cilindričnim cevima različitog poluprečnika. Obično se uzima ista dužina za ove cilindrične segmente koja iznosi 5÷10 mm. Na ovaj način se dobija akustički model vokalnog trakta. Ekvivalentno električno kolo se dobija kaskadnim povezivanjem T-četvoropola kojima se modeluju cilindrični segmenti pojedinačno. Prenosna karakteristika ovako dobijenog ekvivalentnog električnog kola predstavlja prenosnu karakteristiku vokalnog trakta. Maksimumi u ovoj prenosnoj karakteristici predstavljaju formantne frekvencije izgovorene foneme.

Zidove vokalnog trakta grade meka tkiva konačne krutosti. Sa povećanjem pritiska unutar vokalnog trakta dolazi do njegovog "naduvavanja" što se u krajnjoj konsekvenci manifestuje na promenu oblika, odnosno poprečnog preseka vokalnog trakta. U principu, isto važi i za samu komoru maske jer se njene obrazine uglavnom prave od elastičnih materijala.

Na osnovu rezultata eksperimentalnih merenja, sa cilindričnim cevima, dobijena je linearna zavisnost između pritiska unutar cevi i promene njenog poluprečnika:

$$\Delta r = pC \quad (1)$$

gde je: Δr promena poluprečnika cevi, p je pritisak unutar cevi, a C elastičnost zida cevi.

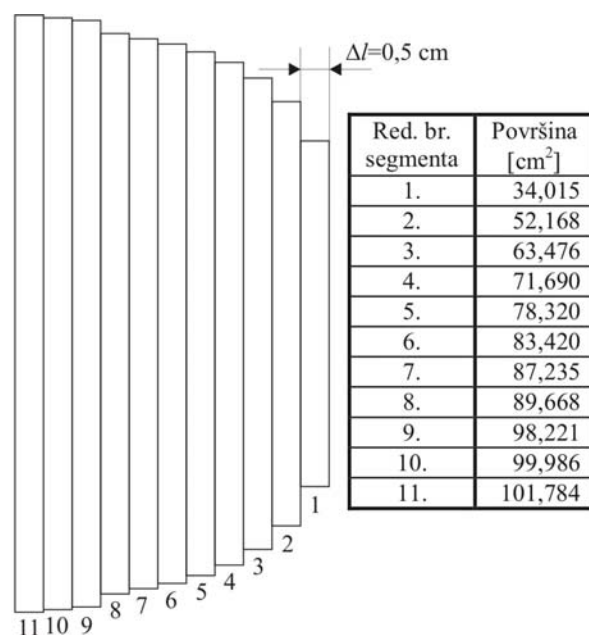
Dakle, ako se u cilindričnoj cevi, poluprečnika r , pritisak poveća za p_{NAD} doći će do povećanja površine poprečnog preseka cevi i ona će iznositi:

$$A_{\text{NAD}} = \pi(r_0 + \Delta r)^2 = \pi(r_0 + p_{\text{NAD}}C)^2 \quad (2)$$

Poluprečnik r_0 je početni poluprečnik cevi, tj. kada natpritisk ima nultu vrednost.

Prema brojnim eksperimentalnim merenjima za elastičnost zida vokalnog trakta [3], [4] može se uzeti prosečna vrednost u iznosu od 2×10^{-4} cm/Pa.

Na bazi relacije (2) izvršena je promena oblika vokalnog trakta u toku izgovora ruskih vokala. Konfiguracije vokalnog trakta u toku izgovora ruskih vokala su preuzete iz [5]. Korišćen je model vokalnog trakta sa gubicima (termalnim i viskozim) [6], a impedansa zračenja otvora usta je modelovana električnim kolom kojim se simulira zračenje klipa postavljenog na sferu [6]. U simulacijama je korišćena maska zapremine 280 cm³ čiji je oblik prikazan na Sl.1. Maska je nezračećeg tipa i sačinjena je od cilindričnih cevi dužine 5 mm površina poprečnih preseka datih u tabeli na Sl.1.



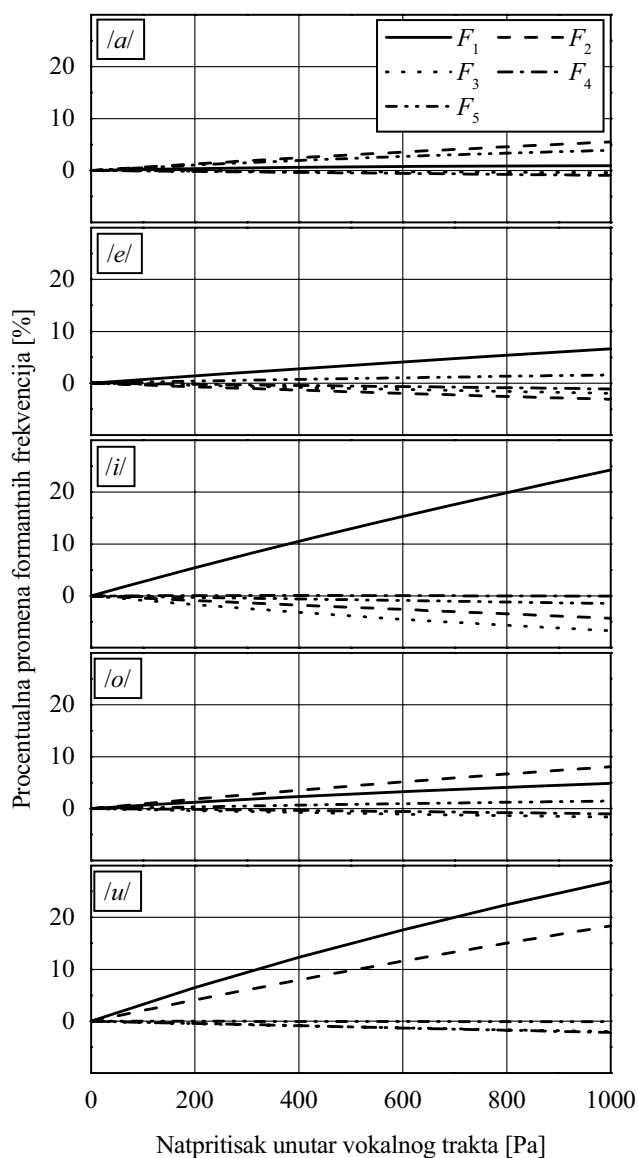
Sl. 1. Maska zapremine 280 cm³ korišćena u simulacijama.

Natpritisk unutar vokalnog trakta je menjan od 0 do 1000 Pa sa korakom 200 Pa. Za svih ovih šest vrednosti natpritiska estimirano je prvih pet formantnih frekvencija ruskih vokala.

III. REZULTATI SIMULACIJA

Procentualne promene formantnih frekvencija vokala u zavisnosti od natpritiska u vokalnom traktu prikazane su na Sl.2. Procentualne vrednosti su računane u odnosu na slučaj kada u vokalnom traktu nema natpritiska.

Uticaj natpritisaka unutar vokalnog trakta na formantne frekvencije vokala je velik i mora mu se posvetiti dužna pažnja. Tako na primer, frekvencija prvog formanta vokala /u/ i /i/ se povećava za više od 20% ako je natpritisak 1000 Pa. Ovo su ogromne promene sa stanovišta sistema za automatsko prepoznavanje govora. Najveće promene formantnih frekvencija su izmerene kod vokala /u/ i /i/, a nešto manje promene (ispod 10%) kod preostala tri vokala. U principu, promene formantnih frekvencija koje izaziva natpritisak su nelinearne. Uglavnom su veće promene kod nižih formanta. Najveće promene su izmerene kod prvog formanta. Promene frekvencija drugog formanta su manje izražene, dok se promene ostalih formanta u potpunosti mogu zanemariti.



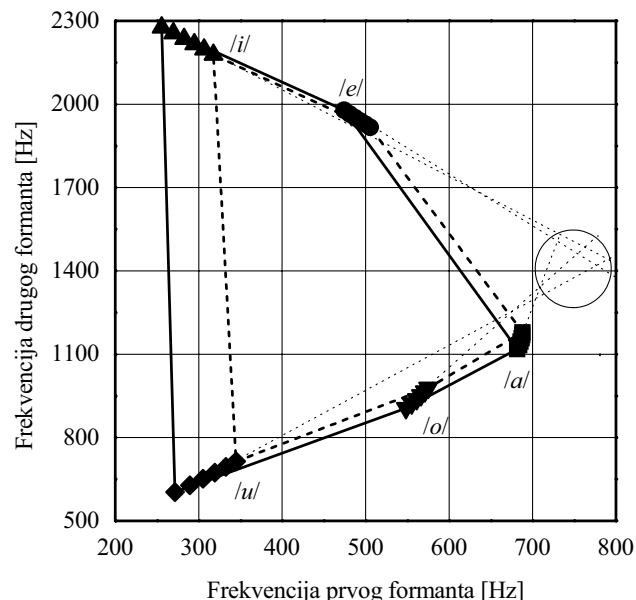
Sl. 2. Procentualne promene formantnih frekvencija ruskih vokala u funkciji natpritisaka unutar vokalnog trakta.

Isto tako, može se zapaziti da su promene većeg intenziteta kod vokala koji imaju konfiguracije vokalnog trakta sa malom površinom poprečnog preseka otvora usta ili dela vokalnog trakta na mestu suženja "jezik-nepce". Površina otvora usta kod izgovora vokala /u/ iznosi 0,65

cm², a isto toliko iznosi i površina poprečnog preseka vokalnog trakta na mestu suženja "jezik-nepce" kod izgovora vokala /i/. Promene površine poprečnog preseka koje izaziva natpritisak prema relaciji (2) procentualno se više ispoljavaju kod užih delova u vokalnog trakta. Zbog toga će konfiguracije vokalnog trakta koje sadrže ovakva uska područja generalno pretrpeti veću promenu oblika, a to znači i da će biti veće promene formantnih frekvencija.

Prema Sl.2 ne može se sa sigurnošću konstatovati trend promena formantnih frekvencija. Jedino što je jasno sa ove slike je da natpritisak najviše utiče na frekvencije prva dva formanta. Kod nekih vokala frekvencija prvog formanta (ili drugog formanta) opada sa porastom natpritisaka, a kod drugih vokala ona raste. Ako se nacrtaju F₁-F₂ dijagram situacija postaje malo jasnija.

Na Sl.3 prikazan je standardni F₁-F₂ dijagram za ruske vokale u funkciji natpritisaka. Svaki pojedinačni simbol predstavlja jednu vrednost natpritisaka koji je menjan od 0 do 1000 Pa sa korakom 200 Pa. Tanke tačkaste linije pokazuju pravac promena formantnih frekvencija vokala kada se povećava natpritisak. Ove linije se "seku" u okolini tačke koja odgovara frekvenciji prvog formanta oko 750 Hz i frekvenciji drugog formanta oko 1400 Hz (obeleženo krugom na Sl.3). Drugačije rečeno, formantne frekvencije svih vokala teže jednoj jedinstvenoj vrednosti, odnosno oblik vokalnog trakta teži jedinstvenom obliku. Zbog postojanja natpritisaka unutar vokalnog trakta dolazi do njegovog "naduvavanja" tako da sve konfiguracije vokalnog trakta počinju da poprimaju sličan oblik bez obzira koji se vokal izgovara. Otuda i približavanje istim formantnim frekvencijama.



Sl. 3. F₁-F₂ dijagram ruskih vokala u funkciji promene natpritisaka unutar vokalnog trakta.

Grubo računato, za cev zatvorenu sa oba kraja dužine 23 cm prva rezonansa je u okolini frekvencije 740 Hz, a druga u okolini 1480 Hz. U analizi koja je ovde urađena korišćena je maska zapremine V=280 cm³ koja ima dubinu 5,5 cm. Ako se uzme da je dužina vokalnog trakta u toku izgovora ruskih vokala 17-18 cm onda to odgovara

gornjem proračunu. Natpritisak unutar vokalnog trakta jednostavno dovodi do toga da vokalni trakt poprima formu uniformne cilindrične cevi zatvorenu na oba kraja.

Prema Sl.3, polje F_1 - F_2 se sužava kada se povećava natpritisak unutar vokalnog trakta (opseg promena formantnih frekvencija nacrtan isprekidanom linijom u poređenju sa opsegom nacrtanim punom linijom). To direktno dovodi do uticaja na tačnost estimacije formantnih frekvencija kod sistema za automatsko prepoznavanje govora. Tačnost estimacije formantnih frekvencija se smanjuje jer su diskriminacione oblasti oko vokala manje.

U prethodnoj analizi uzete su u obzir samo promene oblika vokalnog trakta, a ne i komore maske. Sličnom analizom, može se pokazati da će formantne frekvencije vokala blago pasti ako se menja samo oblik maske zbog prisustva natpritisaka. Promene se uglavnom odnose na frekvencije prva dva formanta i manje su od 1% računato u odnosu na slučaj kada natpritisaka nema. U suštini, frekvencije formanta se približavaju vrednostima koje odgovaraju normalnom govoru, tj. govoru kada se ne koristi kiseonička maska. Prisustvo natpritisaka povećava zapreminu komore maske tako da se smanjuje uticaj maske na formantne frekvencije vokala. U krajnjem slučaju, ako je zapremina maske beskonačno velika to je isto kao da maska nije ni postavljena na lice govornika. Natpritisak na neki način poništava uticaj maske na govor, ali samo u simboličnom iznosu.

Ako se simulacija uradi za istovremeno dejstvo natpritisaka na oblik vokalnog trakta i kiseoničke maske dobiće se rezultati koji se neznatno razlikuju od onih prikazanih na Sl.2 i Sl.3.

Rezultati simulacija prikazani u ovom radu principijelno se slažu sa rezultatima istraživanja drugih autora [7]. I u ovim istraživanjima konstatuje se "centralizacija" formantnih frekvencija uzrokovana prisustvom natpritisaka. U [7] je korišćen model vokalnog trakta bez kiseoničke maske tako da formantne frekvencije teže frekvencijama $F_1 \approx 550$ Hz i $F_2 \approx 1570$ Hz. Ove frekvencije približno odgovaraju prvim dvema četvrttalasnim rezonansama uniformne cilindrične cevi otvorene na jednom kraju i dužine oko 17 cm (prosečna dužina vokalnog trakta kod odraslog muškarca). Kao što je pokazano u ovom radu, kada se u model vokalnog trakta uključi i kiseonička maska onda se njegova "efektivna" dužina povećava za dužinu maske i više nije u obliku cevi otvorene na jednom kraju već u obliku cevi zatvorene na oba kraja.

IV. DISKUSIJA I ZAKLJUČAK

Rezultati obavljenih simulacija pokazuju da disanje sa natpritisakom ima značajne posledice na proces govora, odnosno na formantnu strukturu izgovorene foneme. Promene formantnih frekvencija su posledica, pre svega, promenjenog oblika vokalnog trakta. Grubo rečeno, dolazi do "naduvavanja" vokalnog trakta. Isto se dešava i sa maskom, ali ove promene nemaju bitniji uticaj na formantne frekvencije vokala.

Uticaj natpritisaka se manifestuje najviše na niže

formante: prvi i drugi. Za natpritisak u iznosu od 1000 Pa frekvencije prvog formanta vokala se povećavaju i do 27% u poređenju sa slučajem kada nema natpritisaka. Za isti ovaj iznos natpritisaka frekvencije drugog formanta se menjaju u granicama od -4 do +18%. Promene frekvencija prvog i drugog formanta su takve da dolazi do njihove "centralizacije" jer teže istim vrednostima bez obzira na to koji je vokal izgovoren.

U svakom slučaju, promene formantnih frekvencija izazvane natpritisakom u vokalnog trakta su značajne i moraju se uzeti u obzir kada se projektuju sistemi za upravljanje govorom.

Buduća istraživanja u ovom domenu treba da obuhvate merenja u realnim situacijama, tj. merenja promena formantnih frekvencija vokala u uslovima leta vazduhoplova i sa grupom ispitanika (pilota) ne manjom od deset.

ZAHVALNICA

Rad je delom finansiran sredstvima Ministarstva za nauku, tehnologiju i razvoj Republike Srbije preko projekta br. 148028G.

LITERATURA

- [1] Badin P., Fant G., "Notes on vocal tract computation", *STL-QPSR 2-3/1984*, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, 1984, pp. 53-108.
- [2] Flanagan J., *Speech analysis, synthesis and perception*, New York, Springer-Verlag, 1972.
- [3] Ernesting J., "Some effects of raised intrapulmonary pressure in man", *AGARD Monograph, AGARDOGRAPH 106*, Technivision Ltd., Maidenhead, 1966.
- [4] Svirsky M., Stevens K., Matthies M., Manzella J., Perkell J., Wilhelms-Tricarico R., "Tongue surface displacement during bilabial stops", *J. Acoust. Soc. Am.*, Vol. 102, 1997, pp. 562-571.
- [5] Fant G., *Acoustic theory of speech production*, Mouton, The Hague, 1970.
- [6] Vojnović M., Mijić M., "An improved model for the acoustic radiation impedance of the mouth based on an equivalent electrical network", *Applied Acoustics*, 66, 2005, pp.481-499.
- [7] South A., "A model of vowel production under positive pressure breathing", *Proceedings of EUROSPEECH-2001*, 2001, pp. 1515-1518, Aalborg, Denmark.

ABSTRACT

The speech control systems are based on the speech recognition techniques. Use of the speech control systems are particularly problematic because of many negative factors which have influence on accurate of speech recognition process: high level of ambient noise and vibration, use of oxygen mask, great pilot psychophysical exposure etc. Specific of oxygen mask use is positive pressure breathing. Simulation results shows that breathing with pressure of about 1000 Pa have enormous influence on the first two vowel formants. Range of formant frequency change is more than twenty percentages in compare with case without positive pressure.

INFLUENCE OF THE POSITIVE PRESSURE BREATHING ON THE VOWELS FORMANT FREQUENCIES

Milan Vojnović