INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

Differentiating acoustic and physiological features in speech for hypoxia detection

Benjamin O'Brien[1], *Adrien Gresse*[2], *Jean-Baptiste Billaud*[2], *Guilhem Belda*[2], *Jean-François Bonastre*[1]

[1]Laboratoire Informatique d'Avignon, EA 4128, Avignon Université, France
[2]Semaxone, France

benjamin.o-brien@univ-avignon.fr, adrien.gresse@semaxone.fr

Abstract

In order to stave off the effects of hypoxia, speech may become limited at elevated altitudes. This paper evaluates the role of speech on acoustic and physiological features used to detect hypoxia. Acoustic, cerebral blood oxygenation, and cardiac signals were recorded from participants who completed control and normobaric hypoxia experimental conditions. Acoustic and physiological features were extracted from (non-)speech segments via a voice activity detection method. Support Vector Machines were used to evaluate hypoxia classification using independent and combined features produced at sea-level and simulated 5 km altitudes. Models were built upon a 4-fold cross-validation design and evaluated on an independent dataset. Our results confirmed the importance of physiological features when detecting hypoxia. When combined, acoustic features boosted performance by 10% at 5 km in comparison to sea-level. Hypoxia detection may be improved by distinguishing respiration from non-speech.

Index Terms: hypoxia; speech production; voice activity detection; multi-modal modeling;

1. Introduction

It is well established that hypoxia is a danger to aviation medicine [1, 2]. Hypoxia corresponds to a lack of oxygenation that generally occurs when an individual is at a high altitude. Hypoxia can severely reduce concentration required to complete complex tasks, which can lead to accidents and death. [3] demonstrated the lasting cognitive and psychomotor effects of hypoxia, as reductions in cerebral oxygen saturation via 60 min of exposure to 7.5% of fraction of inspired oxygen (FiO_2) at high altitude (25,000 ft) lead to reduced cognitive functioning. Similar effects were reported in [4], which showed hypoxia not only impaired working memory, but that it also increased the average error rate of responses of participants performing the Paced Auditory Serial Addition Test at a simulated altitude of 10 km. Given the cognitively demanding tasks required of pilots, the goal of the current study was to identify acoustic and physiological features associated with hypoxia.

Numerous studies have shown that various vocal characteristics, such as voice onset time (VOT) [5, 6, 7, 8, 9], fundamental frequency [10], formants [11], and harmonicity [12, 13] are affected by hypoxia. [5] showed VOT separation width and response times decreased at higher altitudes. [8] had participants say non-sense two-syllable words under hypobaric conditions and reported an increase in fundamental frequency range and a decrease VOT. Upon initial exposure to higher altitudes, [11] showed that formant intensity decreased, however, increased after acclimatization. These studies highlight the variety of speech related features affected by hypoxia.

Acoustic cues such as a pilot's slurred speech [14] are first signs of hypoxia, however, they may indicate that psychomotor functioning is already impaired. Not only can stress induced by a decrease in oxygen modify voice parameters [5, 6], but it can also lead to a decrease in speech production. Speech may become limited in an effort to stave off the perceivable effects of high altitudes, which, in turn, may limit the presence of certain acoustic features sensitive to detecting hypoxia. Surprisingly, no findings have been reported on the effects of hypoxia on non-speech acoustics, e.g., respiration, pauses, cough, etc. To fill this gap, the current study examined acoustic differences between speech and non-speech produced in normal and normobaric hypoxic conditions.

Given the evidence that hypoxia impairs cognition and motor control, e.g., vocal production, numerous studies have shown that it affects cerebral and arterial blood ventilation. [15] developed a comprehensive review of cerebral blood flow at high altitude. [16] suggested hypoxia led to changes in the cerebral arterial-to-venous (AV) ratio. Similar findings were reported in [17], which showed an indirect relationship between AV ratio and mean arterial blood pressure. [18] showed that hypoxia reduced arterial oxygen saturation, and numerous studies have shown heart-rate variability is an important index for identifying hypoxia [19]. As speech has been shown to reduce ventilation in comparison to non-speaking [20], it is plausible that speech and non-speech may characterize cerebral and arterial blood oxygenation functions differently, which, in turn, may be useful for identifying pre-hypoxic states.

Although various acoustic and physiological features can be derived from time-series based signals, many questions remain as to which characterize hypoxia. The goal of the current study was to identify certain features that describe normobaric hypoxia states. We hypothesized that physiological as opposed to acoustic features might be more effective at detecting hypoxia. By developing models fitted with separated and combined acoustic and physiological features, the goal was to weigh their contributions independently. Since additional psycho-motor processing is required to speak, speech production in conditions with limited oxygen might "overload" processing, which, in turn, may result in changes to acoustic and physiological features. Thus, it was worth examining any effects of speech on selected features used for detecting hypoxia.

2. Methods

2.1. Data

The current study relied on data collected in [21], which focused on the effects of hypoxia on cerebral brain oxygenation. The following briefly describes the participants and experimental protocols developed in the previous study.

10.21437/Interspeech.2023-1678

2.1.1. Participants

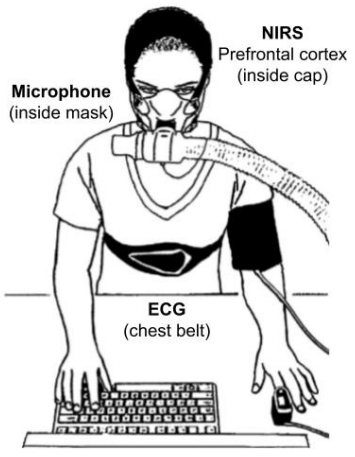
20 students from the Université de Montpellier (France) participated in the study (4 F; mean age: 25 $\pm$ 5 years). Their mean weight and height was 70.1 $\pm$ 8.6 kg and 175.9 $\pm$ 6.2 cm, respectively. All students completed 8 $\pm$ 4 hours per week in University-related sports activities, including handball and swimming. The study was approved by the Ethics Committee of Université de Montpellier.

2.1.2. Experimental protocol

All participants completed two experimental conditions (total duration per condition: approximately 25 min), where the order was randomly selected (non-repeating). Across conditions participants completed an ascension phase from sea-level to 5 km (750 m/s; approximately 6 min 30 sec in order to reach 5 km). The duration at 5 km was 11.8 $\pm$ 0.02 min. Throughout the "normoxia" condition (N), participants were exposed to 21% FiO_2, i.e., at sea-level and 5 km. On the contrary, in the normobaric "hypoxia" (H) condition, participants at sea-level were presented with 21% FiO_2 and at 5 km they were presented with 11.30% FiO_2 via an AltiTrainer device. Thus, for control purposes, during the normoxia condition, participants at 5 km experienced "fake" hypoxia. Only during the hypoxia condition at 5 km did participants experience real hypoxia. All participants completed the two experimental conditions over the span of two to six days. They completed a cognitive task described in [21], however, their performance was excluded from this analysis.

Figure 1 illustrates the technical setup. Speech and respiration were captured with a Shure MX150 microphone placed inside the mask of the AltiTrainer (audio interface: Zoom H6; sampling rate: 44.1 kHz; bit depth: 16). Near-infrared spectroscopy (NIRS) optical sensors (Portalite, AMS) were positioned on the left prefrontal cortex to measure cerebral oxygenation (sampling rate: 10 Hz). Heart rate was monitored via the Polar H10 ECG chest-belt (sampling rate: 1000 Hz).

Figure 1: *Technical setup: microphone (in mask); NIRS sensors (in cap); and ECG (chest belt).*



A custom MATLAB script was written to synchronize acoustic, NIRS, and ECG signals. The computer dedicated to monitoring the ECG signal received digital NIRS signals produced from the Portalite device of Artinis via Bluetooth. To signify the start of data collection, the word "Top" was recorded at the same time as a trigger generated in the Artinis software.

2.2. Feature extractions

2.2.1. Acoustic features

All speech recordings were first down-sampled to 16 kHz. The "Voice Activity Detection" recipe of SpeechBrain [22] was used to segment the audio recordings into "speech" and "non-speech". The time points of each segment were used to separate *a posteriori* ECG and NIRS signals into speech and non-speech segments. 19 MFCCs and the first and second derivatives were extracted and concatenated into 57 feature vectors (window: 1 s; step-size: 0.25 s) and used for acoustic modeling (see Section 2.4). The decision to use a large window and step size was based on the hypothesis that they might be more suitable for characterizing respiration-related sounds rather than short term features more equipped to capture phonetic information.

2.2.2. NIRS-based features

Total hemogoblin (tHb) and relative oxygen saturation ($rSaO_2$) metrics were obtained using conventional methods. First recordings were converted into the Shared Near Infrared Spectroscopy format with the Python package MNE [23]. Raw data was then converted into optical density (OD), whereupon the Temporal Derivative Distribution Repair method [24] was used to correct for recording artifacts. Following this processing the corrected OD data was converted into relative hemoglobin concentration using the Modified Beer-Lambert law. The data were filtered using a bandpass filter (lb: 0.05 Hz; ub: 0.7 Hz). tHb is the sum of HbO_2 (oxyhemoglobin) and HbR (deoxyhemoglobin) and $rSaO_2$ is the ratio between HbO_2 and tHb.

2.2.3. ECG-based features

A custom MATLAB script was developed to identify ECG-based features[1]. Prior to segmentation, ECG signals were filtered using the $designfilt$ bandpass function (lb: 0.5 Hz; ub: 150 Hz). Following the speech segmentation procedures addressed above, R, Q, S, T start, peak, and end points were identified using the $findpeaks$ function. In addition the RR-interval metric was calculated [19]. For each segment, the mean, median, maximum, minimum, and standard deviation was calculated for R, Q, S, T, and RR-interval (25 total features).

2.3. Data processing

2.3.1. Training and testing datasets

Acoustic and physiological features were only extracted during the initial sea-level and simulated 5 km altitudes across experimental conditions. Table 1 describes the number of segments, mean duration per segment (in s), and total duration across altitudes, experimental conditions, and speech types.

During the data collection, some participants were unable to complete both experimental conditions. As such, the data were partitioned into *training* and *testing* datasets. The training dataset was composed of 9 participants who completed both conditions (N: 9782.8 s; H: 9710.0 s), whereas the testing dataset was composed of 11 participants who completed one condition (N: 9609.3 s; H: 2075.7 s).

2.3.2. Selecting of physiological features

Stepwise linear discriminant analysis (SLDA) was used to select which physiological features were sensitive to distinguish-

[1] Scripts are available at: https://github.com/OBrienBenjamin/ECG_Features

Table 1: *Number of segments (n), mean duration per segment (s), and total duration of segments across altitudes (SL: sea-level; 5km: 5 km), experimental conditions (Normoxia, Hypoxia), and speech types (S: speech; NS: non-speech).*

		Normoxia			**Hypoxia**		
		n	mean (s)	total (s)	n	mean (s)	total (s)
SL	S	175	3.7 ± 2.6	650.6	124	4.4 ± 5.3	547.2
	NS	188	37.7 ± 51.6	7125.9	134	31.4 ± 52.3	4208.1
5km	S	210	3.5 ± 2.3	740.9	142	3.6 ± 2.9	515.9
	NS	211	51.5 ± 58.7	10874.6	141	46.2 ± 57.6	6514.4

ing hypoxie and normoxia conditions in the training dataset. Only segments extracted at simulated 5 km altitudes across conditions were used. The R function $greedy.wl$ (in the $klaR$ R-package) was used, as it implements a forward SLDA using the Wilks' Lambda criterion. The forward SLDA method was preferred, as it starts from the null hypothesis and incrementally adds new variables with the highest discriminant power based on Wilks' Lambda value until $p > 0.05$. Table 2 illustrates the physiological features selected from the SLDA using Wilks' Lambda criterion.

Table 2: *SLDA for selecting physiological features sensitive to classifying training dataset conditions using the Wilks' Lambda criterion. {**, ***} represent $p < \{0.01, 0.001\}$, respectively.*

Feature	**Statistic**	**Wilks' Lambda**	F	p
RR-interval	Median	0.70	191.55	***
$rSaO_2$	Median	0.65	37.21	***
R peak	Mean	0.63	12.70	***
Q peak	Median	0.6	20.82	***
S peak	Median	0.59	9.37	**
Q peak	Std. Dev	0.58	9.00	**

2.4. Model training

2.4.1. Balancing data

The number of segments were not consistent across the two conditions, and, due to segments not having the same duration, the number of sliding frames were not equal. Thus, as a way of obtaining an equal distribution across experimental conditions, random sampling with replacement procedures were use to select the same number of frames for training and testing datasets.

2.4.2. Cross-validation

The classification task was performed on acoustic and physiological features extracted from training dataset (non-)speech segments. First data normalization procedures were performed according to the training dataset mean and standard deviation. Models were then fitted with acoustic, physiological, or combined features using the Python scikit-learn library. In order to determine which model configuration was the most accurate distinguishing experimental conditions, a 4-fold cross-validation design was performed on the training data with a splitting ratio set to 0.15 for validation. A Support Vector Machines (SVM) model was selected over Logistic Regression and Stochastic Gradient Descent models, as the goal was to identify the best margin between hypoxia and normoxia conditions, as opposed to optimizing decision boundaries based solely on the testing dataset. This decision was based on the limited size of the training dataset, which was balanced, as extracted features were derived from nine participants who completed both experimental conditions. Each model configuration was fitted and validated on the four folds 10 times, whereupon their results were averaged. The best model configuration relied used a linear SVM with default regularization coefficient preceded by a Nystroem mapping of 2048 components (RBF kernel; $\gamma = 0.75$).

2.5. Evaluation

Models fitted to the training dataset were evaluated with the testing dataset. The same pre-processing procedures were applied to the testing dataset with the exception of normalization statistics computed on the training dataset in order to removed bias. Model predictions were performed on 1 s sliding frames, which were then reduced to a single prediction per segment by selecting the most frequent class predicted for every frames of the segment. All model predictions were compared against ground-true labels. We ran 10 iterations of the whole train-test procedure and global accuracy was reported as X $\pm$ Y%, where X and Y are means and standard deviations, respectively.

3. Results

3.1. Model training results

Table 3 and Figure 2 show hypoxia classification accuracy results by 4-fold cross-validation design models fitted with features extracted from the training dataset (non-)speech segments across altitudes. In general, model predictions were enhanced when distinguishing hypoxia and normoxie conditions at 5 km in comparison to sea-level. Interestingly, acoustic models were less accurate when evaluating speech segments at either altitude. Due to the limited number of training dataset speakers, it is possible that the models became sensitive to speaker-dependent speech patterns and failed to make accurate predictions on new, different speakers. Performance was boosted when combining features, with stronger performances observed at both altitudes when models evaluated non-speech segments.

Table 3: *Hypoxia detection accuracy of 4-fold cross-validation of models fitted with features extracted from training dataset across altitudes (SL: sea-level, 5km: 5 km) and speech types (S: speech; NS: non-speech).*

		Model		
		acoustic	physiological	combined
SL	S	$72.2 \pm 2.1\%$	$82.6 \pm 7.9\%$	$86.2 \pm 2.7\%$
	NS	$84.8 \pm 4.2\%$	$81.2 \pm 4.6\%$	$92.8 \pm 2.6\%$
5km	S	$71.5 \pm 3.0\%$	$90.3 \pm 2.9\%$	$88.2 \pm 2.3\%$
	NS	$89.5 \pm 2.0\%$	$88.9 \pm 2.0\%$	$96.0 \pm 1.7\%$

3.2. Evaluation results

Models fitted on the training dataset were evaluated with the testing dataset. The following describes the hypoxia detection performance of the combined features model that evaluated all testing dataset segments across altitudes: 40% hit rate; 29% correct rejection rate; 6% missed alarm rate; 18% false alarm rate; and 7% indecisive due to segment durations less than 1 s.

Table 4 shows the hypoxia classification accuracy (%) results of models fitted with acoustic, physiological, and com-

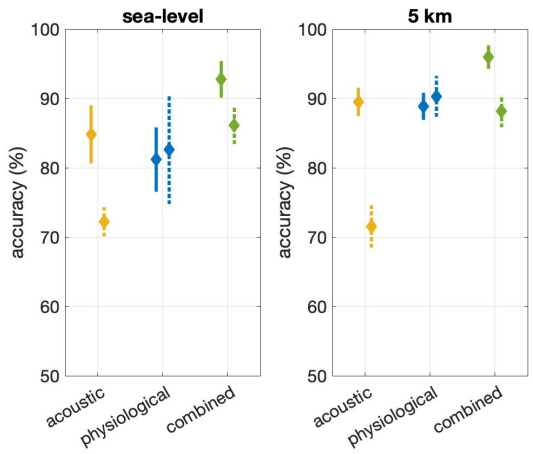


Figure 2: *Hypoxia classification accuracy on models fitted to features extracted from training dataset. Diamonds represent means, while solid and dotted lines represent standard deviations of non-speech and speech segments, respectively.*

bined features extracted from the testing dataset (non-)speech segments across altitudes (sea-level, 5 km). In general there was an overall drop in performance when evaluating the (unbalanced) testing dataset. Although acoustic models were performing at chance levels, an approximate 10% performance boost was observed when the combined features model evaluated either speech type at 5 km altitude. Because oxygen levels were not reduced at sea-level altitudes (FiO_2: 21%), these results suggest that the combined features models were more accurate at 5 km altitudes due to changes in oxygen level. Considering the physiological response of speech, a 5% gap between speech types at 5 km were observed, which suggests speech production induced more effort. This observation signals a relationship between speech production and physiological variables due to hypoxia, which could explains why classification performance was increased with the combined features models.

Table 4: *Hypoxia detection accuracy of models fitted with features extracted from testing dataset across altitudes (SL: sea-level, 5km: 5 km) and speech types (S: speech; NS: non-speech).*

		Model		
		acoustic	physiological	combined
SL	S	51.3%	60.3%	63.6%
	NS	52.7%	61.3%	62.6%
5km	S	52.0%	65.3%	71.6%
	NS	54.9%	60.9%	72.4%

4. Discussion

The goal of the current study was to examine the effects of speech on acoustic and physiological features used for detecting hypoxia. Following the selection of cerebral and arterial blood oxygenation features, our model training with a 4-fold cross validation design revealed strong results, where the combined features models yielded a mean hypoxia detection accuracy of approximately 90% for training dataset (non-)speech segments across altitudes. However, when evaluating these models with the testing dataset, performance dropped approximately 20%. While the quality and size of the datasets likely contributed to this decline in performance, our findings offer insight on the effects of hypoxia with regards to the influence of speech on features used for automatic detection of hypoxia.

Our SLDA results revealed two major findings regarding the physiological features sensitive to detecting hypoxia at 5 km. First, they support findings suggesting that hypoxia affects certain cerebral ($rSaO_2$[16, 17]) and cardiac (RR-interval, R-peak[19]) features. The selected physiological features in combination enhanced model performance when discriminating between (non-)speech segments produced at 5 km across experimental conditions. Second, the inclusion of Q and S peak features suggests additional cardiac signal characteristics might be valuable when detecting hypoxia. This observation could prove useful by expanding the scope of blood oxygenation features sensitive to detecting normobaric hypoxia.

Our training results revealed that acoustic models slightly enhanced performance in comparison to physiological models when evaluating non-speech segments at sea-level (+4%) and 5 km (+1%) altitudes. Interestingly, a mean drop in performance (-15%) was observed when acoustic models evaluated speech segments across altitudes. Given the limited size of the dataset, models may have learned speaker-dependent speech characteristics, which, when absent in other speakers, lead to a decrease in performance. To confirm this explanation, the development of a larger dataset is in progress, where, following an initial voice activity detection segmentation, the goal is to distinguish "respiration" sounds embedded in non-speech segments.

Our evaluation results revealed that our combined features models had a mean boost of 10% when evaluating (non-)speech segments at 5 km in comparison to sea-level altitudes. While only physiological feature models were capable of discriminating experimental conditions without net distinction over speech types, when acoustic information was added models outperformed feature-independent models. By providing both acoustic and physiological information, the models developed relations between feature types, which suggests relevancy in adding new modalities for tasks such as hypoxia detection. This may be tied to physiological changes implicated with speech production, which, in a hypoxic state, may become disturbed in a compensatory manner. Because speech production involves the physiological function of respiration, which has been linked to cardiac information, it is possible that by combining acoustic and physiological features the model considers respiration information. Our findings suggest that a multimodal model may improve hypoxia detection performance.

5. Conclusion

This study examined the effects of speech on acoustic and physiological features used for detecting hypoxia. The following lists take-away messages from the study: (1) the physiological features selected to characterize hypoxia were consistent and expanded upon existing literature; (2) the combined acoustic and physiological feature models performed at a relatively high accuracy level; and (3) model performance dropped when evaluating a testing dataset, suggesting dependence on participants.

6. Acknowledgment

This work is associated to the Skyphisia project 21015563 admitted by Region Occitanie for the Readynov2021 grant in deliberation on the CP/2021-OCT/10.04.

7. References

[1] G. PC and P. P, "In-flight hypoxia - still a worrying bane," *Indian Journal of Aerospace Medicine*, vol. 54, 2010.

[2] C. Neuhaus and J. Hinkelbein, "Cognitive responses to hypobaric hypoxia: Implications for aviation training," *Psychology Research and Behavior Management*, vol. 7, pp. 297–302, 11 2014.

[3] K.-T. Chiang, M.-Y. Tu, C.-C. Cheng, H.-H. Chen, W.-W. Huang, Y.-L. Chiu, Y.-Y. Wang, and C.-Y. Lai, "Contributions of hypoxia-awareness training to the familiarization of personal symptoms for occupational safety in the flight environment," *International Journal of Environmental Research and Public Health*, vol. 18, no. 6, 2021. [Online]. Available: https://www.mdpi.com/1660-4601/18/6/2904

[4] C. Malle, P. Quinette, M. Laisney, C. Bourrilhon, J. Boissin, B. Desgranges, F. Eustache, and C. Piérard, "Working memory impairment in pilots exposed to acute hypobaric hypoxia," *Aviation, space, and environmental medicine*, vol. 84, pp. 773–9, 08 2013.

[5] P. Lieberman, A. Protopapas, E. Reed, J. W. Youngs, and B. G. Kanki, "Cognitive defects at altitude," *Nature*, vol. 372, no. 6504, pp. 325–325, Nov. 1994. [Online]. Available: http://www.nature.com/articles/372325a0

[6] Lieberman, Philip, Morey, Angie, Hochstadt, Jesse, Larson, Mara, and Mather, Sandra, "Mount Everest: A Space Analogue for Speech Monitoring of Cognitive Deficits and Stress," *Aviation, Space, and Environmental Medicine*, 2005.

[7] M. Van Puyvelde, W. Vanderlinden, M. Van den Bossche, T. Bucovaz, T. De Winne, X. Neyt, and N. Pattyn, "When the hypoxia "silent killer" starts to talk: The early detection of pre-symptomatic hypoxia through voice stress analysis," *Journal of Science and Medicine in Sport*, vol. 20, p. S44, Nov. 2017. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S144024401731191X

[8] M. Van Puyvelde, X. Neyt, W. Vanderlinden, M. Van den Bossche, T. Bucovaz, T. De Winne, and N. Pattyn, "Voice Reactivity as a Response to Acute Hypobaric Hypoxia at High Altitude," *Aerospace Medicine and Human Performance*, vol. 91, no. 6, pp. 471–478, Jun. 2020. [Online]. Available: https://www.ingentaconnect.com/content/10.3357/AMHP.5390.2020

[9] M. Van Puyvelde, E. Debecker, X. Neyt, F. Detaille, W. Vanderlinden, and N. Pattyn, "Flying High. Voice Stress Analysis to Detect Pre-symptomatic Acute Hypobaric Hypoxia at 25000 Ft," in *Advances in Safety Management and Human Performance*, P. M. Arezes and R. L. Boring, Eds. Cham: Springer International Publishing, 2021, vol. 262, pp. 97–104, series Title: Lecture Notes in Networks and Systems. [Online]. Available: https://link.springer.com/10.1007/978-3-030-80288-2_12

[10] Khan, M, Sondhi, S, Vijay, R, Sharma, SK, and Salhan, AK, "Fundamental Frequency of Voice under Normobaric and Hypobaric Hypoxia," *Indian Journal of Aeropsace Medicine*, 2015.

[11] S. Sondhi, M. Khan, R. Vijay, A. K. Salhan, and S. K. Sharma, "Effect of Normobaric and Hypobaric Hypoxia on Formant Characteristics of Human Voice," *International Journal of Computer Applications*, vol. 122, no. 15, pp. 32–37, Jul. 2015. [Online]. Available: http://research.ijcaonline.org/volume122/number15/pxc3905055.pdf

[12] Z. N. Milivojevic, M. Milivojevic, and D. Brodic, "The Effects of the Acute Hypoxia to the Fundamental Frequency of the Speech Signal," *Advances in Electrical and Computer Engineering*, vol. 12, no. 2, pp. 57–60, 2012. [Online]. Available: http://www.aece.ro/abstractplus.php?year=2012&number=2&article=10

[13] Z. N. Milivojevi, D. Brodi, and M. Z. Milivojevi, "The Effects of the Active Hypoxia to the Speech Signal Inharmonicity," vol. 23, no. 2, p. 6, 2014.

[14] B. J. Johnston, G. S. Iremonger, S. Hunt, and E. Beattie, "Hypoxia training: symptom replication in experienced military aircrew," *Aviat. Space Environ. Med.*, vol. 83, no. 10, pp. 962–967, Oct. 2012.

[15] P. Ainslie and A. Subudhi, "Cerebral blood flow at high altitude," *High altitude medicine & biology*, vol. 15, pp. 133–40, 06 2014.

[16] F. Wong, T. Alexiou, T. Samarasinghe, V. Brodecky, and A. Walker, "Cerebral arterial and venous contributions to tissue oxygenation index measured using spatially resolved spectroscopy in newborn lambs," *Anesthesiology*, vol. 113, pp. 1385–91, 11 2010.

[17] M. Bo Rasmussen, V. Eriksen, B. Andresen, S. Hyttel-Sorensen, and G. Greisen, "Quantifying cerebral hypoxia by near-infrared spectroscopy tissue oximetry: The role of arterial-to-venous blood volume ratio," *Journal of Biomedical Optics*, vol. 22, p. 025001, 02 2017.

[18] L. B. Rowell, D. G. Johnson, P. B. Chase, K. A. Comess, and D. R. Seals, "Hypoxemia raises muscle sympathetic activity but not norepinephrine in resting humans," *J. Appl. Physiol.*, vol. 66, no. 4, pp. 1736–1743, Apr. 1989.

[19] J. Coppel, P. Hennis, E. Gilbert-Kawai, and M. P. Grocott, "The physiological effects of hypobaric hypoxia versus normobaric hypoxia: a systematic review of crossover trials," *Extrem. Physiol. Med.*, vol. 4, no. 1, p. 2, Feb. 2015.

[20] S. E. Baker, J. Hipp, and H. Alessio, "Ventilation and speech characteristics during submaximal aerobic exercise," *J. Speech Lang. Hear. Res.*, vol. 51, no. 5, pp. 1203–1214, Oct. 2008.

[21] E. Fresnel, G. Dray, S. Pla, G. Belda, and S. Perrey, "Cerebral oxygenation and cognitive performance under acute normobaric hypoxia," in *ACAPS 2021 - 19ème congrès international des chercheurs en Activités Physiques et Sportives*, 2021.

[22] M. Ravanelli, T. Parcollet, P. Plantinga, A. Rouhe, S. Cornell, L. Lugosch, C. Subakan, N. Dawalatabad, A. Heba, J. Zhong, J.-C. Chou, S.-L. Yeh, S.-W. Fu, C.-F. Liao, E. Rastorgueva, F. Grondin, W. Aris, H. Na, Y. Gao, R. De Mori, and Y. Bengio, "SpeechBrain: A general-purpose speech toolkit," 2021.

[23] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. S. Hämäläinen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013.

[24] F. A. Fishburn, R. S. Ludlum, C. J. Vaidya, and A. V. Medvedev, "Temporal derivative distribution repair (tddr): A motion correction method for fnirs," *NeuroImage*, vol. 184, pp. 171–179, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811918308103