ICSV13 - Vienna
The Thirteenth International Congress on Sound and Vibration
Vienna, Austria, July 2-6, 2006

SUBJECTIVE SPEECH INTELLIGIBILITY MEASUREMENTS IN A LIGHT AIRCRAFT CABIN UNDER LABORATORY CONDITIONS

Tino Bucak[1], Marko Horvat[2], and Doris Novak[3]

[1]Department of Aeronautics, Faculty of Transport and Traffic Sciences
Vukelićeva 4, 10000 Zagreb, Croatia
[2]Department of Electroacoustics, Faculty of Electrical Engineering and Computing
Unska 3, 10000 Zagreb, Croatia
[3]Croatian Air Force
tino.bucak@fpz.hr

Abstract
The quality of speech communications inside the aircraft cabin is one of the primary issues in aircraft safety. After comprehensive field investigations of light aircraft cabin noise during flight and the results of speech intelligibility obtained, attempt was made to create simulated cabin noise field inside the hull of a flight training device by using cabin noise audio clips digitally prerecorded in real environment. A group of qualified subjects was put under subjective speech intelligibility tests and the most significant results are presented.

INTRODUCTION

The interfering influence of interior (*cabin* and/or *cockpit*) aircraft noise on speech communication has been a severe problem for a number of years already, at least since engine-driven aircraft became common means of transportation. The problem features particularly G/A aircraft, both fixed and rotary wing, i.e. light piston engine propeller-driven aircraft and helicopters, but this is the case in turboprop aircraft as well, especially light single-engine trainers, utility transporters and small commuters. The propellers, helicopter rotors, poorly damped exhaust systems located fairly near the cockpit and acoustically porous fuselage are factors which result in rather noisy aircraft interiors [1].

An especially adverse characteristic of cabin noise is its similarity to the typical speech spectrum, which implies that cockpit noise is a very efficient speech masker. The masking theory is based on the knowledge of physiology of basilar membrane, and implicates that the increase in the hearing threshold of a sound, due to the concurrent presence of some other (masking) sound is the result of psychoacoustic jam, i.e. functional inaccessibility of cochlear receptors due to the *priority rule.*

The most successful speech masking is achieved by the masking sound in the frequency band of 300 Hz to 500 Hz, in which the cockpit noise components are also dominant, and due to the so-called *spreading effect* the masking of the entire speech spectrum is possible. This affects most the consonants, the carriers of speech information, the levels of which are generally low while the frequency components in the higher spectrum range. Therefore, the masking of speech by noise is in fact the problem of masking the consonants [2].

Although objective methods of measuring speech intelligibility can rather simply provide reliable results, the testing of speech intelligibility by subjective methods inside of small G/A aircraft during flight is very complex, since the testing procedure assumes engagement of a larger number of subjects-listeners in strictly controlled acoustic conditions, which is technically and organizationally difficult to achieve. Therefore, the real conditions have been simulated by means of synthetic trainer device (STD) cabin, where it was possible to organize several consecutive tests on a sufficient number of subjects.

SUBJECTIVE METHODS OF TESTING SPEECH INTELLIGIBILITY

Syllable intelligibility

Testing of syllable intelligibility is one of the oldest subjective methods of measuring the intelligibility and quality of speech communication. It was presented in 1929 by Fletcher and Steinberg as a method for assessing the quality of the speech communication channel by means of its articulation values. In their testing the articulation was defined as the percentage of transmitted speech units that were correctly received by the listener, and the speech units could be phonemes, monosyllables, words or sentences. It is obvious that their methodology allows forming of two types of testing: *Intelligibility Test* or, more precisely, the test of understanding, and *Articulation Test* or *Syllable Articulation Test* [3].

The intelligibility tests are carried out by meaningful words and sentences, and the results are evaluated by the percentage of correct definition of the meaning of what has been said. The results of testing the intelligibility of words and sentences will be expectedly higher than the results of articulation tests, since the listener may correct the inaccurately received words and sentences taking their overall meaning into consideration by means of the guessing method using one's own redundancy. On the other hand, the articulation tests or syllable articulation tests are usually done with simple voice stimuli that have no specific meaning in a language, commonly known as nonsense monosyllables. Depending on the language, the following combinations are possible: CVC (consonant-vowel-consonant) and CCV (consonant-consonant-vowel). The test proceeds in such a way that one person reads the in-advance prepared text consisting of a series of nonsense monosyllables and on the receiving side a number of persons listen and record what they hear. The intelligibility of nonsense monosyllables is defined as a mean value in the percentage of correctly received nonsense monosyllables in relation to the total number of nonsense monosyllables uttered. The paper presents the results of speech intelligibility tests carried out on a group of thirty subjects by means of tables of nonsense monosyllables derived from Croatian language.

The intelligibility of nonsense monosyllables of 96% is considered equivalent to natural speech, intelligibility from 85% to 95% is very good, from 75% to 85% good, from 65% to 75% the speech may yet be understood but with additional effort, whereas the intelligibility of less than 65% is unsatisfactory. The intelligibility in ideal conditions is not the expected 100% due to the influence of the psychological-linguistic factors (garbled articulation, excessive speech rate, stress, pitch and intonation of speech, etc.).

Other subjective methods

Most of other subjective methods of measuring speech intelligibility are mainly variations of the nonsense monosyllable tests and tests of meaningful words, such as e.g. RT – *Rhyme Tests* and the resulting MRT – *Modified Rhyme Tests* and the DRT – *Diagnostic Rhyme Tests*, as well as PBW – *Phonetically Balanced Words*, etc. Finally, the simplest straight method is the so-called CJM – *Category Judgement Method*, i.e. evaluation (on a scale from 1 to 5) of the quality of the listened-through speech, as well as the quality of the connection between two speakers (e.g. conversation via two-way radios) etc. The latter method is usually applied in everyday flying practice during pre-flight check of the quality of radio communication between the aircraft crew and the air traffic controller (so-called *radio-check*).

THE EXPERIMENT

Measurement procedures and equipment layout

The original aircraft noise was recorded by the portable PC within the cabin of the aircraft *Cessna* 172R in real conditions. The *wav* file was then processed and set into closed loop, thus providing continuous source of cabin noise. In processing the signals, the start and the end of the audio sequence were overlapped with excessive care in amplitude and phase synchronization, in order to avoid signal distortion, dropouts and odd sounds like "cracks" and "clicks" that might influence the concentration of the subjects during audio testing.

Spectra Lab® ver. 4.32. software program was used for digital recording and FFT analysis of noise, while Cool Edit Pro® ver. 2.0. software program served for digital processing. Laboratory testing of speech intelligibility in adverse conditions (high levels of aircraft cabin noise) was performed at the flight simulation laboratory of the Faculty of Transport and Traffic Sciences in Zagreb, inside the hull of synthetic flight trainer device *BT Austria Flight Simulators* #BT-220 (Figure 1) on a qualified sample of volunteers from the group of students of the aeronautical course attending classes for military and civil pilots as well as air traffic controllers. The examinees passed a complete medical check at the Institute for Aviation Medicine at the Clinical Hospital Dubrava in Zagreb and have clear audiometric findings. They were included in the tests by random selection method.

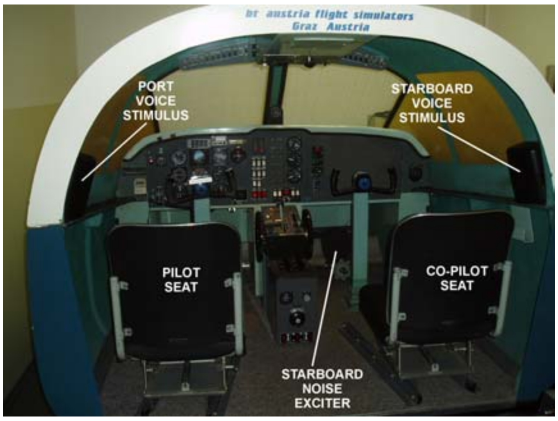

Figure 1 – General layout of loudspeakers for reproduction of speech and cabin noise within BT-220 flight training device cabin

During speech intelligibility tests the subjects were exposed to cabin noise ranging from 65 dBA to 85 dBA in consecutive increments of 5dB and two different levels of voice stimuli, 62dB (which corresponds to average level of normal speech) and 75dB (which corresponds to average level of loud speech).

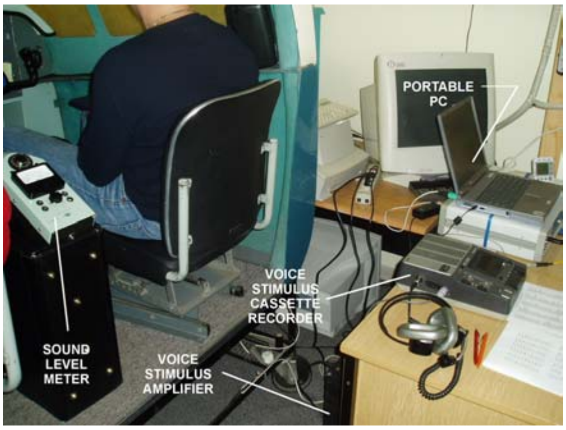

Figure 2 – Measurement set for subjective testing of speech intelligibility

The voice stimulus signal was recorded by the portable cassette-recorder *Philips AAC 4000 Language Trainer*, otherwise used for speaking drills at the laboratory for radiotelephony (Figure 2). Six groups, each containing ten double CVC/CCV nonsense monosyllable combinations were recorded. Since the physiologic-articulation factor in the range of voice levels necessary for the planned audio tests (62 – 75 dB) practically does not affect the intelligibility [4], for the recording of the voice stimulus only one speaker was used, whose speech was of constant intonation and level, and the necessary voice stimuli levels in reproduction were obtained with adequate volume gain. The aircraft cabin noise signal (*noise excitation*) was transmitted from the portable PC to the power (*noise exciter*) amplifier, with loudspeakers located at the bottom of the cockpit in front of the control pedals, and the voice stimulus signal was transmitted through the communication channel (interphone) of the simulator to the loudspeakers in the cockpit (Figure 3). The cockpit noise level was adjusted through the software application Cool Edit Pro® installed on a portable PC and controlled by Sound Level Meter centered inside the cockpit.

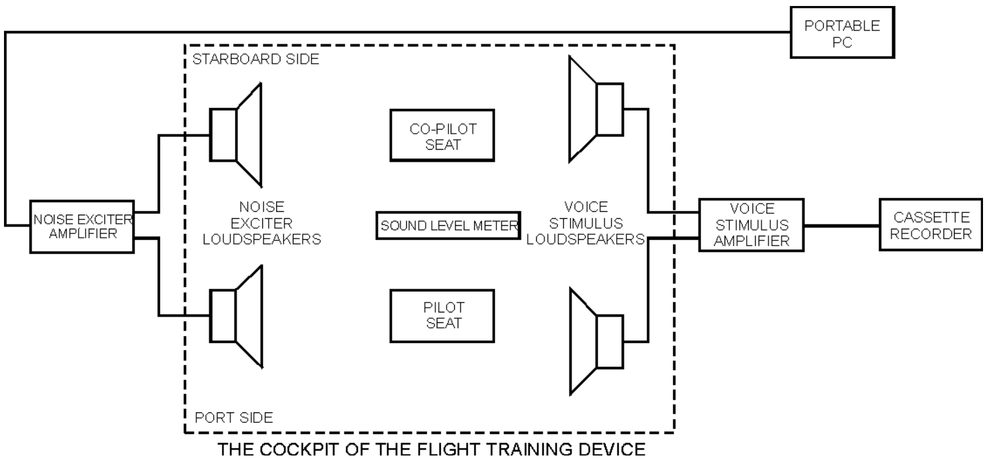

Figure 3 – Schematic of the set for subjective testing of speech intelligibility in laboratory conditions

Testing consisted of two different groups of tests, CVC and CCV nonsense monosyllables, with two levels of voice stimuli and five levels of aircraft noise, with two subjects undergoing the test simultaneously, who sat on the pilot and co-pilot seat. The subjects listened to the recorded speech and noted down what they heard (or thought they did) into the form prepared in advance (Figure 4).

Figure 4 – The procedure of subjective speech intelligibility testing in laboratory conditions

The review of the results

After having performed the testing, the recorded nonsense monosyllables were compared to the original tables, and based on the ratio between correct and incorrect ones the intelligibility was calculated in percentages for every subject. To make the calculation more straightforward, the psychological-linguistic factor has been neglected, which otherwise slightly affects the final result.

The obtained results of testing the intelligibility of nonsense monosyllables (CVC, CCV and the combination CVC/CCV) for different A-weighted cockpit noise levels and two different speech levels have been processed statistically and presented in Table 1 and Figures 5 and 6. The table shows also the standard deviation, whose value, as expected, rises with the increase of cockpit noise level.

Table 1 – The results of speech intelligibility test using nonsense monosyllables

CABIN NOISE LEVEL [dBA]	THE INTELLIGIBILITY OF THE NONSENSE MONOSYLLABLES [%]											
	CVC				CCV				CVC/CCV			
	normal speech		loud speech		normal speech		loud speech		normal speech		loud speech	
	$\overline{X}$	σ	$\overline{X}$	σ	$\overline{X}$	σ	$\overline{X}$	σ	$\overline{X}$	σ	$\overline{X}$	σ
65	81	17,6	91	10,5	90	8,7	96	10,1	86	9,82	93,3	6,1
70	61	12,7	88	6,8	81,1	16,9	83,3	10	71	11, 7	86	7,7
75	49	19,7	68	18,6	59	19	86	11,3	54	16,5	77	13,2
80	40	24,5	66	19,4	44,4	21,9	68	19,2	42,2	22,5	67	16,2
85	23	29,6	50	25,5	20	23,3	56	16,7	21,3	25,5	53	20,2

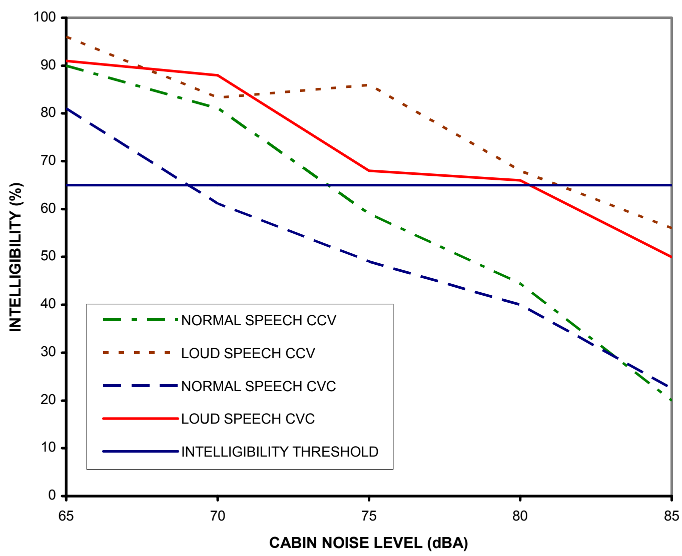

Figure 5 - Dependence of intelligibility of CVC and CCV nonsense monosyllables on A-weighted levels of cabin noise

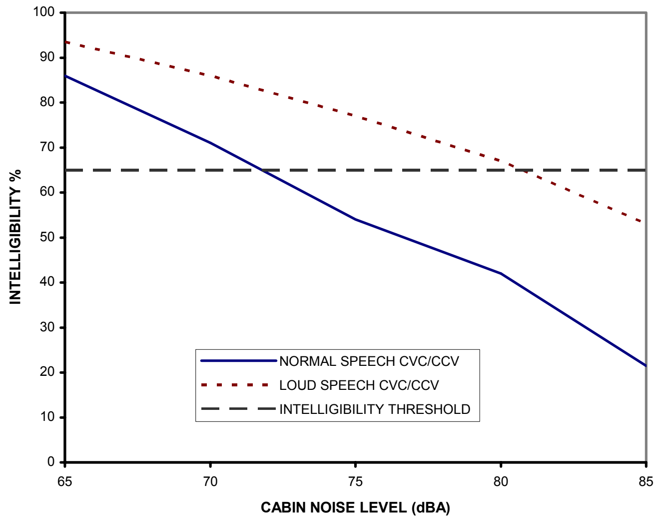

Figure 6 – Dependence of combined CVC/CCV nonsense monosyllables intelligibility on A-weighted cabin noise level

Mathematical modelling of results using Statistica® software application has provided approximation functions of correlation between A-weighted noise levels L_A and percentage intelligibility of CVC/CCV combination of nonsense monosyllables for normal (I_{ns}) and loud speech (I_{ls})

$$I_{ns} = -3{,}2L_A + 292 \qquad [\%] \tag{1}$$

$$I_{ls} = -2L_A + 225 \qquad [\%] \tag{2}$$

which could enable, with acceptable accuracy, fast and simple calculation of percentage intelligibility of nonsense monosyllables derived from Croatian language, based on the measured A-weighted light aircraft interior noise.

CONCLUSION

Subjective methods of measuring speech intelligibility require a larger number of subjects and strictly controlled acoustic conditions to achieve valid results. Therefore, it is rather difficult to organize tests of spoken information intelligibility by subjective methods within the cockpit of a light G/A aircraft in real in-flight conditions. In order to perform the planned testing, real conditions were simulated in the flight simulation laboratory by means of the synthetic trainer device cockpit, where it was possible to arrange a number of consecutive tests on a qualified sample of thirty subjects randomly selected.

Nonsense monosyllable tables derived from the Croatian language were used for voice stimuli. They were uttered by a single speaker, pre-recorded on a cassette-recorder and then reproduced to the subjects at two different levels, normal (62dB) and loud (75dB) while masked by five different levels of cockpit noise (65 – 85 dBA in 5 dB increments) pre-recorded *in-situ* and adequately post-processed. The results have shown almost linear dependence of the nonsense monosyllable intelligibility on the cabin noise level and have confirmed recommended interior noise levels in order to maintain the acoustical comfort and, essential for the flight safety, reliable speech communication. More research in this field continues by using voice stimuli with meaningful words typical for aviation communication.

REFERENCES

[1] Smith, M. J. T.: *Aircraft Noise*, Cambridge University Press, 1989.

[2] Shepherd, W. T.: *Cockpit Speech Interference Considerations*, Proceedings of the 3rd Symposium on Aviation Psychology, Ohio State University, 1985, p. 411-417.

[3] Quackenbush, S. R.; Barnwell, T. P.; Clements, M. A.: *Objective Measures of Speech Quality*, Prentice Hall, New Jersey, 1988.

[4] Pickett, J. M.: *Limits of Direct Speech Communications in Noise*, J. Acoust. Soc. Am., Vol. 30, No. 4, 1958.