

Effects of simulated helicopter cabin noise on intelligibility and annoyance

MALCOLM D. ARNOULT, JAMES W. VOORHEES, and LYNNE G. GILFILLAN
Texas Christian University, Fort Worth, Texas

Helicopter cabin noise was simulated by combining a broadband signal (pink noise, or PN) with a triad of pure tones (PT) at 650, 1900, and 5000 Hz. Each component was presented at four loudness levels (0, 60, 70, and 80 dB(A)), with all 16 combinations arranged in two unsystematic orders. Intelligibility was measured by means of sentences to be judged as true or false. A male speaker presented 10 sentences at each noise condition. One group of subjects heard the sentences at 50 dB(A) and another group at 55 dB(A). The annoyance of each noise combination was rated by the same subjects on a 5-point scale. Both sources and their interaction were significantly related to reduced intelligibility and increased annoyance at both speech levels. In general, the PN component was more responsible for loss of intelligibility, and the PT component more responsible for annoyance.

Investigations of the effects of aircraft noise on human beings have been devoted primarily to effects on listeners who are outside of the aircraft rather than to effects on passengers (Clevenson & Leatherwood, 1979). In the case of conventional aircraft, either propeller or jet, the differences between the interior and exterior sounds are not great, and both kinds of sound can be adequately simulated by a broadband signal such as pink noise. In the case of helicopters, however, there are substantial differences between the sound signals, primarily due to the fact that for cabin noise, in addition to the broadband sound, there are conspicuous components that approximate pure tones.

Investigations of the behavioral effects of helicopter cabin noise have potential commercial value, because the high levels of cabin noise that currently characterize helicopters tend to discourage their use as business aircraft. Since the procedures that would attenuate the broadband and pure tone components are different (Garinther & Hodge, 1981), it would be useful to know the effects of each. Clevenson and Leatherwood (1979) investigated this problem by selective filtering, first, of a 2700-Hz component, then of a 1370-Hz component, and, finally, of all frequencies above 500 Hz. They found that reduction in the pure tone components improved intelligibility of spoken words and reduced the rated annoyance of the noise. Because they were selectively filtering recorded noises, however, removal of each pure tone reduced the overall loudness levels of the stimulus, thereby making it difficult to determine how much of the change in be-

havioral effects was due to the removal of the tones alone. In the present study, the helicopter cabin noise was simulated by combining pink noise with a triad of pure tones (650, 1900, and 5000 Hz) so that the two components of the noise could be varied independently.

Also, the present study investigated simultaneously both the annoyance caused by the aircraft noises and the effects on speech intelligibility. Rupf (1977) asked raters to judge the acceptability of simulated cabin noises of general aviation aircraft as backgrounds for conversation, including estimates of the length of time they would be willing to conduct a conversation at each loudness level. Key and Powell (1980) related external helicopter noise levels to ratings of conversational ease on scales such as "difficulty of conversation" and "increased vocal effort." They also found that, at the highest noise level of 80 dB(A), 33% of the subjects stopped conversing. None of these measures, however, contributed significantly to the determination of the annoyance measure. Unlike these previous investigations, the present study was able to compare annoyance ratings with actual interference of communication rather than with ratings of estimates of such interference.

METHOD

Helicopter Noise

The single-frequency components of the simulated helicopter noise were pure tones (PT) at 650, 1900, and 5000 Hz. These tones were generated by audio oscillators and combined in a Sony MX-650 mixer. The combined tones were set equal in power by means of a Crown RTA-2 Analyzer and fed into one channel of a Yamaha CA-800 stereo amplifier.

The broadband component was pink noise (PN) generated by the Crown RTA-2 and fed into the other channel of the Yamaha CA-800. These two acoustic components (PT and PN) were separately connected to two Braun (Output C) speakers located 61 cm apart at a distance of 1.83 m behind and 46 cm above the head of the subject.

Each of the acoustic components could be presented at any of four loudness levels, 0, 60, 70, or 80 dB(A), measured at the point of the subject's ears. Each component was set at its nominal loudness individu-

This research was supported by Grant No. NAS 2-69 from NASA Ames Research Center, Moffett Field, CA. A preliminary investigation was supported by a grant from Bell Helicopter Textron, Fort Worth, TX. J. W. Voorhees's present address is: MS 239-21, NASA Ames Research Center, Moffett Field, CA 94035. L. G. Gilfillan's present address is: Lynne Gilfillan Associates, Inc., 3900 Tamworth Road, Fort Worth, TX 76116. Requests for reprints should be sent to M. D. Arnoult, Department of Psychology, Box 32878, Texas Christian University, Ft. Worth, TX 76129.

ally; consequently, the combined output was slightly higher than either single component. The four loudness values of the two components were combined into 16 separate stimulus conditions. Two quasirandom sequences of these loudness conditions were used.

Verbal Stimuli

Investigators of intelligibility have used a variety of spoken materials, such as monosyllabic words, phonemes, phonemically balanced words, rhyming words, and complete sentences. Since this research was concerned with spoken discourse, complete sentences were used. Hudgins, Hawkins, Karlin, and Stevens (1947) developed an audition test based on a series of simple questions to which one-word answers could be given. The materials developed for the present research followed their approach, and many of the needed sentences were produced by converting their questions into simple declarative sentences that were clearly either true or false. Additional items were developed from lists of phonemically balanced words that were utilized at critical points in new sentences. Altogether, 160 sentences were developed.

These sentences were recorded on audio tape by a male speaker in 16 sets of 10 sentences each, with each set having a total duration of approximately 50 sec. Within each set of 10 sentences, 5 were true and 5 were false, with the order systematically varied from set to set. The sets were also equated with respect to difficulty level and phonemic properties.

The recorded sentences were transferred to one channel of an audio cassette and played through a speaker located 1.83 m in front of the subject at eye level. The other channel of the cassette contained 400-Hz beeps that were heard only by the experimenter and used to control the timing of the experimental procedures. One group of subjects ($N=20$) heard the sentences at an average loudness of 55 dB(A), and the other group of subjects ($N=16$) heard the sentences at an average loudness of 50 dB(A).

Subjects

The subjects were 14 males and 22 females ranging in age from 18 to 40. All were students who received academic credit for volunteering to participate.

Procedure

The experiment was conducted in an acoustically shielded, moderately anechoic chamber. At the subject's position was a console containing eight small signal lights and seven pushbutton switches. One group, consisting of three lights and two switches, was used for responding to the sentences. A central blue light marked "X" was illuminated whenever a sentence was being spoken, whether the speech was audible or not. At the completion of each sentence, red and green lights marked "T" and "F" were illuminated, and the subject was required to push one of the associated pushbutton switches to indicate whether the sentence was true or false.

Each noise condition started about 7 sec before the first sentence was spoken and continued for 2 sec after the last sentence was completed. At this point, there was a 10-sec period of silence during which the other five lights on the console were illuminated. These lights, numbered 1 to 5, corresponded to the five positions on a printed scale of annoyance. The descriptive labels associated with the scale positions were as follows: 1, not annoying; 2, slightly annoying; 3, moderately annoying; 4, very annoying; 5, extremely annoying. During this 10-sec period of silence between noise conditions, the subject was required to push the switch associated with the appropriate annoyance judgment.

The true-false judgments and the annoyance judgments were indicated to the experimenter by means of signal lights and were manually recorded. In addition, a permanent record was established by means of a printing recorder.

The instructions to the subject emphasized that there was no trickery associated with the true-false statements and that the only source of difficulty in the task would be the noise conditions. It was also emphasized that a response must be made whenever the red and green lights came on, even if the subject could not understand the sentence or did not hear it all. In the event that any problem developed, the subject could communicate with the experimenter by means of an open intercom.

RESULTS

Intelligibility

Responses to the ten true-false statements in each of the noise segments were scored by the formula "right minus wrong plus 10," so that a score of 20 represented errorless performance and a score of 10 represented chance performance. Mean intelligibility scores for the 16 noise conditions are shown in Table 1 for the 50-dB(A) and 55-dB(A) speech levels. A 4 (PN) \times 4 (PT) \times 2 (order of presentation) analysis of variance for the 50-dB(A) speech level showed highly significant ($p < .001$) main effects for both PT and PN and for the PT \times PN interaction. In addition, the order variable was also significant ($p = .047$), but no interactions involving order reached significance. The analysis of the data obtained at the 55-dB(A) speech level also showed significant main effects for PT and PN ($p < .001$) and for their interaction, but the order variable was not significant.

Annoyance

Mean annoyance ratings for the 16 noise conditions are shown in Table 2 for the 50-dB(A) and 55-dB(A) speech levels. Analyses of variance performed on both sets of data showed, as was the case with intelligibility, highly significant main effects for PT and PN and the interaction between them ($p < .001$ in each case). For these data, the order variable was significant in both analyses ($p = .003$), but no interactions involving order were significant.

DISCUSSION

Examination of Table 1 shows that a clear effect of adding pure tones (PT) to the pink noise (PN) occurred only at the PN level of 60 dB(A). The failure of the PT variable to show any systematic effect on intelligibility at the 0-dB(A) and 80-dB(A) levels of PN is understandable in that those performances are at the upper and lower limits, respectively, that are possible for the task. It is not clear, however, why the addition of PT appears to have no systematic effect on intelligibility when PN is at 70 dB(A). At that level, performance is not constrained by the scoring limits, but there are no significant differences in performance related to differences in PT loudness. It can also be observed that this pattern of relationships exists at both levels of speech loudness, although

Table 1
Mean Intelligibility Scores as a Function of Stimulus Loudness

Pink Noise (PN) in dB(A)	Pure Tones (PT) in dB(A)			
	0	60	70	80
Speech Level = 50 dB(A)				
0	19.4 (1.2)	19.5 (0.9)	19.4 (1.0)	19.2 (1.2)
60	19.4 (1.0)	18.6 (1.7)	17.3 (2.7)	13.9 (2.3)
70	11.0 (3.4)	12.4 (3.6)	11.4 (2.7)	11.6 (2.2)
80	9.3 (2.6)	9.6 (2.4)	9.1 (1.9)	9.1 (3.1)
Speech Level = 55 dB(A)				
0	19.6 (0.9)	19.1 (1.0)	19.6 (1.0)	19.2 (1.3)
60	19.6 (1.2)	19.6 (1.0)	18.9 (1.5)	16.9 (1.8)
70	15.4 (4.7)	14.2 (3.2)	15.1 (2.7)	13.8 (4.4)
80	9.0 (2.4)	10.0 (3.5)	9.8 (2.7)	10.4 (2.7)

Note—Standard deviations are in parentheses.

Table 2
Mean Annoyance Ratings as a Function of Stimulus Loudness

Pink Noise (PN) in dB(A)	Pure Tones (PT) in dB(A)			
	0	60	70	80
Speech Level = 50 dB(A)				
0	1.05 (0.22)	2.19 (0.54)	2.75 (0.93)	3.88 (1.15)
60	2.13 (0.72)	2.31 (0.79)	2.62 (0.80)	3.69 (0.87)
70	3.50 (0.89)	3.50 (0.94)	3.81 (1.11)	4.19 (0.66)
80	4.30 (0.70)	4.25 (0.77)	4.31 (0.87)	4.67 (0.49)
Speech Level = 55 dB(A)				
0	1.0 (0.00)	2.20 (1.02)	2.95 (1.00)	3.90 (1.12)
60	1.70 (0.47)	1.80 (0.55)	2.40 (0.75)	3.50 (0.89)
70	3.20 (0.41)	3.20 (0.81)	2.90 (0.94)	3.90 (0.91)
80	4.30 (0.67)	4.40 (0.60)	4.60 (0.60)	4.70 (0.57)

Note—Standard deviations are in parentheses.

all of the effects of noise are somewhat attenuated at the higher speech level of 55 dB(A).

The results of the annoyance ratings are more complex. When PN = 0 dB(A), annoyance increases with increasing loudness of PT, even though intelligibility is unaffected. As the loudness of PN increases, the effect of PT on annoyance becomes progressively less apparent, disappearing completely at 80 dB(A). As was the case with the intelligibility data, similar results were obtained at both loudness levels of the speech, except that the effects of PT were less apparent at the 70-dB(A) level of PN.

Overall, these results are similar to the results obtained by Clevenson and Leatherwood (1979), but there also appear to be important differences. They found that removing the PT components from recorded helicopter noises improved intelligibility, but those results are complicated by the fact that the removal of the PT components also reduced the overall loudness level of the stimulus. They concluded, however, that "the small region of overlap between the two curves (73–76 dB[A]) indicates an improvement in listening task performance of approximately 15 percent for the case where tones are removed" (p. 8). It may be noted, however, that the region of overlap between the two noises (73–76 dB[A]) is very close to the loudness of their speech stimuli (76 dB[A]). In the present experiment, the loudness of the PT component is strongly related to intelligibility when the loudness level of the PN component is close to the loudness level of the speech.

With respect to annoyance ratings, Clevenson and Leatherwood (1979) found little or no effect of removing PT components beyond effects attributable to an overall loudness reduction. The present experiment, on the other hand, showed a clear effect of PT on annoyance, except at the highest level of pink noise. Perhaps the difference is that, with more loudness levels of each type of noise independently presented, the interactive effects of PN and PT on annoyance were more clearly revealed.

Finally, the kind of interaction between these two components that appeared in this experiment is discouraging with respect to the selection of procedures for improving the acoustic environment of the helicopter cabin. Clearly, it is the broadband component of the noise that most interferes with intelligibility and which should be attenuated in order to improve conversation. At the same time, at low levels of PN, annoyance is strongly related to the loudness of the PT component. Improvement in intelligibility, therefore, will not necessarily lead to a reduction in annoyance.

REFERENCES

- CLEVENSON, S. A., & LEATHERWOOD, J. D. (1979). *Effect of helicopter noise on passenger annoyance* (NASA Technical Memorandum 80106). Hampton, VA: NASA Langley Research Center.
- GARINER, G. R., & HODGE, D. C. (1981). *The background and bases for the proposed military standard on acoustical noise limits in helicopters* (Technical Memorandum 5-81). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.
- HUDGINS, C. V., HAWKINS, J. E., KARLIN, J. E., & STEVENS, S. S. (1947). The development of recorded auditory tests for measuring hearing loss for speech. *Laryngoscope*, *47*, 57-89.
- KEY, K. F., & POWELL, C. A. (1980). *Effects of conversation interference on annoyance due to aircraft noise* (NASA Technical Paper 1712). Hampton, VA: NASA Langley Research Center.
- RUFF, J. A. (1977). *Noise effects on passenger communication in light aircraft*. (Publication 770446). Warrendale, PA: Society of Automotive Engineers, Inc.

(Manuscript received for publication November 12, 1985.)