Tangney
2313AS

Final Technical Report

AFOSR Grant No. F49620-93-1-0070

Auditory Perception

Marion F. Cohen
Department of Communication Sciences
The University of Connecticut
Storrs, CT 06269

05-18-00P04:43 RCVD

REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

0290

38

Public reporting burden for this collection of information is estimated to average 1 hour per respor
gathering and maintaining the data needed, and completing and reviewing the collection of inform
collection of information, including suggestions for reducing this burden, to Washington Headquar
Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Bu

j data sources,
r aspect of this
1215 Jefferson
20503.

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE

3. Final - 01 Jan 93 - 31 Oct 97

4. TITLE AND SUBTITLE
Auditory Perception

5. FUNDING NUMBERS
F49620-93-1-0070

6. AUTHOR(S)
Dr Marion F Cohen

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Connecticut
Storrs CT 06269

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NL
801 N. Randolph St.,Rm 732
Arlington VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release: Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The experiments described were designed to measure how well the acoustic cues which are generally understood to be used for sound source segregation are processed by listeners who have more than the usual difficulty with that task. It is generally agreed that certain acoustic characteristics of competing signals are used by the auditory system for signal separation. These include coherence of amplitude modulation and coherence of frequency change across frequency, as well as similarity of a preceding signal with a delayed signal. We report here normative date for these tasks, which were collected on trained young subjects with normal auditory sensitivity. Based on the data produced by our experiments on signal separation, we conducted further experiments which verify our assumptions that these experimental paradigms do, in fact, reflect the signal separation abilities of the auditory system. Specifically, we have collected data on subjects with normal sensitivity who exhibit an unusual amount of difficulty listening in noise. The results of these experiments should be of interest in determining which of the several experimental paradigms thought to be related to signal separation are indeed necessary to performing that task.

14. SUBJECT TERMS
Acoustic, Auditory System

15. NUMBER OF PAGES
10

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
UNCLASS	UNCLASS	UNCLASS	

Standard Form 298 (Rev. 2-89) (EG)
Prescribed by ANSI Std. 239.18
Designed using Perform Pro, WHS/DIOR, Oct 94

Marion F. Cohen

Grant No. AFOSR F49620-93-1-0070

ABSTRACT

The experiments described were designed to measure how well the acoustic cues which are generally understood to be used for sound source segregation are processed by listeners who have more than the usual difficulty with that task. It is generally agreed that certain acoustic characteristics of competing signals are used by the auditory system for signal separation. These include coherence of amplitude modulation and coherence of frequency change across frequency, as well as similarity of a preceding signal with a delayed signal. We report here normative date for these tasks, which were collected on trained young subjects with normal auditory sensitivity. Based on the data produced by our experiments on signal separation, we conducted further experiments which verify our assumptions that these experimental paradigms do, in fact, reflect the signal separation abilities of the auditory system. Specifically, we have collected data on subjects with normal sensitivity who exhibit an unusual amount of difficulty listening in noise. The results of these experiments should be of interest in determining which of the several experimental paradigms thought to be related to signal separation are indeed necessary to performing that task.

This grant, which had been awarded for a three year period, was continued into the fifth year by virtue of a no-cost extension. In the first three years, we completed essentially all of the proposed experiments designed to study sound source segregation, with an emphasis on monaural echo suppression.

Our research the first year was directed towards determining the specific acoustic cues used by the auditory system to fuse echoes with the previously received primary sound. This suppression of multiple hearings is similar in several respects to sound source segregation, in that the auditory system must first recognize that the echo is acoustically similar to the primary sound, before it can be suppressed. For the auditory system to recognize a stimulus as an echo, it must meet two criteria: it must be delayed relative to a previously received stimulus, and it must be acoustically similar to that stimulus. The experiments are based on the assumption that when a signal is perceived as an echo, its detectability will be reduced.

Experimental series I:

Our first series of experiments was performed to demonstrate that a signal, intended to simulate an echo, is more difficult to detect when it is acoustically identical to the masker, intended to simulate an echo. Two acoustic cues, frequency similarity and harmonicity, are examined.

Experiment 1: The masker was a 500 Hz sinusoid presented at 70 dB SPL. The signal, also a sinusoid, was either 500 or 750 Hz and was presented either before or after the masker. Results show that signal detectability improves with increasing temporal shift between signal and masker, regardless of the direction of that shift. When the signal and masker are identical, detectability of the leading signal improves quickly as a function of increasing temporal shift, but improvement in detectability of the delayed signal is considerably slower, resulting in marked asymmetry in detection thresholds of the leading and delayed signal. It is noteworthy that for signal delays up to 30 ms, the average signal detectability remains nearly constant, despite the fact that as much as 30 ms of the signal is presented with no masker. This would not be predicted by existing temporal masking data. For the 750 Hz masker, the data are nearly symmetrical. That is, detectability of the leading and delayed signals are almost identical. Thus, the data indicate that an acoustically identical signal and masker results in considerably more masking for a delayed signal than the leading signal, but when the signal and masker are different in frequency, we do not observe that asymmetry.

Experiment 2: It is possible, in examining the above data, that the effectiveness of the masker is responsible for the observed asymmetry. The following control experiment was performed to determine if that is the case. In that experiment, the masker was a narrow band noise, selected to be as effective as the 500 Hz masker but without being acoustically similar to the 500 Hz signal. The masker was a noise band (100 to 1000 Hz) with an overall level of 70 dB SPL. The signal was a 500 Hz sinusoid. The data show that for the noise band masker, the asymmetry observed for the 500 Hz sinusoidal masker is nearly eliminated for two of the three subjects and is significantly reduced for the third subject When the signal and masker are simultaneous, the signal is around 8 dB less detectable with the noise masker than the 500 Hz tone, despite the fact that the noise masker contains considerably less energy in the critical band surrounding the signal. Interestingly, the poorer signal detectability with the noise masker does not persist when the signal is delayed. Rather, the delayed signal becomes more detectable with the noise masker than with the tonal masker. For the leading signal, detectability is not very different for the two maskers, resulting in a more symmetric curve for the noise band masker. These data indicate the asymmetry observed in Experiment 1 cannot be attributed to masker effectiveness.

Experiment 3: This experiment was performed to examine the effect of a harmonic relationship between the delayed signal and the masker on signal detectability. The masker consisted of 3 equal-level sinusoids, 400, 600, and 800 Hz, with an overall level of 70 dB SPL. The signal always had two components which were either 1) included in the masker (400 & 800 Hz), 2) not included in the masker but harmonically related (500 & 700 Hz) or 3) not included and also nonharmonic with the masker (474 & 711 Hz). Results indicate that the inclusive signal is considerably more difficult to detect when it is delayed than when it is leading, followed by the harmonic signal and the nonharmonic signal in that order. These data indicate that a harmonic. relationship between the signal and masker adversely affects detectability of the delayed signal.

Experimental series II:

These experiments were designed to determine the usefulness of harmonicity in providing an acoustic connection between a masker and a signal. The information gained from these experiments will be necessary to design future experiments using a non-simultaneous masking paradigm. The basic assumption was that when harmonic fusion occurs among stimulus components, detection of one of those components will be decreased. The first experiment was designed to test that assumption by comparing detectability of a harmonic and nonharmonic signal in the presence of a harmonic masker. Further, considering the possibility that harmonicity not only fuses the components having the same fundamental, but also separates them from components having other fundamental frequencies, the second experiment was designed to quantify harmonic fusion and segregation respectively.

Experiment 1: These experiments were performed to determine if the existence of an harmonic relationship between the delayed signal and the masker will result in suppression of that signal, as measured by its detectability. Previous research (McFadden, 1988; O'Neill, 1991; Moore, et al 1985a, 1985b; Darwin and Gardner, 1986) has indicated that simultaneous stimulus components are more likely to fuse if they are harmonically related. We first investigated the effect on signal detectability of harmonicity between the signal and a simultaneous masker. The masker, presented at an overall level of 70 dB SPL, consisted of 3 consecutive harmonics of 120 Hz in several frequency ranges. The signal was either harmonically related to the masker or dissonant with it (Plomp & Levelt, 1965). Results show that the harmonic signal is less detectable than the neighboring nonharmonic signal, despite the fact that the harmonic signal is spectrally further from the masker. These data indicate that the effect of harmonicity on signal detectability is relatively independent of spectral masking. A further study was performed to determine if these data might be influenced by the existence of combination tones, which could produce masking on the harmonic signal. A 3-tone nonharmonic masker, with frequency components selected so as to generate a cubic difference tone at 1080 Hz, was used to mask a

1080 Hz signal and an 1100 Hz signal. Results indicate that, on average, the signal at the frequency of the difference ton (1080 Hz) is not more difficult to detect than the 1100 Hz signal. In fact, when compared to the data of the first experiment, the detectability of the 1080 Hz signal is improved, and detectability of the 1100 Hz signal is decreased.

Experiment 2: This experiment was performed to determine if the poorer detectability measured for the signal harmonically related to the simultaneous masker would occur for a non-simultaneous presentation. The masker was 3 equal-level harmonics of 120 Hz (360, 600, and 840 Hz) with starting phases of 0°, -120°, and 45° respectively. The signal was either harmonically related (480 and 720 Hz) or not (516 and 774 Hz). Both signal and masker had a duration of 125 ms including a 10 ms $\cos^2$ rise/fall time. The signal either preceded the masker onset by 10 ms or was delayed by that amount. The masker had an overall level of 70 dB SPL. Results show that both the harmonic and nonharmonic signals are less detectable when delayed relative to the masker than when leading the masker. On average, the delayed harmonic signal is around 11 dB less detectable than the leading signal, and the delayed nonharmonic signal, around 5 dB less detectable. This difference is mainly due to the difference in detectability of the delayed signal.

Several additional experiments were performed to determine the frequency region where harmonicity can affect detection of a delayed signal, and also to determine if the underlying acoustic cue could be pitch similarity of the signal and masker. The results were somewhat variable with change in signals and subjects, but seem to indicate a temporal order effect based on harmonicity, and also, that the pitch of the signal and masker may play a role.

Experimental Series 3:

The alternating (ALT) paradigm was developed to eliminate the perception of echoic color existing in the stimulus of the previous experiments using the leading-delay (LD) paradigm. In the ALT paradigm, there is never a portion of the signal that is, that is, not overlapping in time with any portion of the masker. Also, the structure of the stimulus is such that there are two components of both the signal and masker, and these are alternated, as shown below.

Experiment 1: Segment A in the masker consisted of two equal-level harmonic components, 660 and 880 Hz. Segment B consisted of either a single component or two equal-level harmonic components both with variable frequency. The two segments in the signal were identical to the two corresponding segments in the masker (i.e., segment A' had the same composition as segment A, and segment B' the same as segment B). The signal was either leading or delayed by 50 ms relative the second segment of the masker. The masker was presented at an overall level of 75 dB SPL. Results show that, for every delayed and leading stimulus pair, the delayed signal is less detectable than the leading signal. This detection asymmetry varies with the spectral distance between segments A and B. The asymmetry is minimum when the frequencies of segments A and segment B are close and maximum when the frequencies of segment B are distant from segment A. However, beyond a certain spectral separation, the asymmetry doesn't further increase, while detectability of each leading and delayed signal pair is still improved. The data also show that pitch difference (or fundamental frequency) between two alternated segments affects the amount of detection asymmetry. More asymmetry, as much as 10 dB, is observed when two alternated segments have different pitch. This experiment was performed to study whether the ear is able to use vowel formant structure to identify the echoes of a primary sound. If so, we expect the delayed signal having formants matched to the primary sound would be perceptually suppressed, showing poorer detectability, as compared to either the unmatched delayed signal or the matched leading signal. Three synthetic vowels were used /a/, /I/, and /o/. Each vowel was composed of four formants with each containing two or three adjacent harmonic sinusoids.

Experiment 2: An alternating (ALT) paradigm, was used for this study. The masker consisted of 4 segments, produced by alternating two vowels twice. The four segments (vowel 1-vowel 2-vowel 1-vowel 2) had durations of 100, 125, 125 and 100 ms. The signal had 2 segments (vowel 1-vowel 2 or vowel 2-vowel1), each with a duration of 125 ms. All segments had a 15 ms rise/fall time. The masker was presented at an overall level of 70 dB SPL. Each segment in the masker had the first, third and fourth harmonics of one vowel, with the second formant missing, and each segment in the signal contained only a second formant of a vowel. The entire signal either preceded the second segment of the masker (leading) or was delayed relative to that segment by 25 ms or 50 ms. The detection difference between the delayed and leading signals was termed the "temporal order effect." This was measure when the formant relationship between the corresponding segments of the signal and masker was either matched or unmatched. That is, they would complete a vowel's formant structure if the were matched. The data show that, in general, detection of the delayed matched signal is poorer than that of the leading signal. This is the case for all subjects and both delays. On average, the temporal order effect is about 5 dB for both temporal shifts. Interestingly, for the unmatched vowel configuration (that is, the 2nd formant of the signal does nor make a complete vowel with the masker), the delayed signal is not less detectable than the leading signal, and is, in fact, slightly more detectable.

Experimental Series IV:

Experiment 1: Temporal order effect on detectability of a non-identical signal
Detection asymmetry is observed when the two segments in the signal are identical in frequency to the masker. To examine whether this "frequency identical" is a necessary condition, non-identical signals were used in the same paradigm. For simplification, each segment in the masker and the signal contained one frequency component, and all components were harmonics of 220 Hz. The two segments in the signal had the same frequency distance with the corresponding segments in the masker, and this distance was variable. Other conditions were the same as Experiment 1. Results show that maximum asymmetry (20-30 dB) is obtained when the signal components are identical to the masker's. As frequency difference between the signal and masker increases, the temporal order effect becomes smaller. Peripheral spectral relation between the signal and the masker affects the temporal order effect.

Experiment 2: Temporal order effect in terms of frequency order association
To understand the mechanism underlying this temporal order effect, one basic question is whether this effect is based on the masking level at signal onset. In Experiments 1 and 2, the spectral distance between segments A and A' (or between B and B') is always closer than the distance between segments B and A' (or A and B'). So, when the signal is delayed relative to the second segment of the masker, it starts with a period more masked. Whereas,

when the signal is leading, a easier detected signal period is first evoked. It is interesting to see whether this onset difference would influence signal detection, resulting in detection asymmetry.

Experiment 3: The masker was a harmonic series with missing fundamental (480, 600, 720 Hz; 1320, 1440 1560 Hz; or 3480, 3600, 3720 Hz) with an overall level of 70 dB SPL. The signal, varying in frequency, was a tone either harmonically related to the masker or not but in the same frequency range. The signal and masker were 125 ms in duration and were presented simultaneously. Everything else was as described above. Results indicate that for the low-frequency masker, a harmonic signal is less detectable than the neighboring nonharmonic signal, suggesting that the harmonic relationship between signal and masker caused poorer signal detectability. For the mid-frequency masker, the difference in detectability is even greater, in some instances as high as 30 dB. This persists when the signal is as much as five harmonics away from the masker, which well exceeds the critical bandwidth in that frequency range. For the high-frequency masker, the harmonic signalis less detectable than the nonharmonic signal only when it is one or two harmonics below the masker. Overall, this harmonic fusion seems to be relatively independent of energetic masking. This detection difference between a harmonic and nonharmonic signal could be due to either the harmonic signal fusing with the masker, the nonharmonic signal segregating from the masker, or both. It would be due only to harmonic fusion if the harmonic signal fuses with the masker, causing poorer detectability, while the nonharmonic signal is not fused or segregated and so is detected by a general detection model. Or, it would be due only to harmonic segregation if the nonharmoinic signal is segregated from the masker, resulting in improved detectability, while the harmonic signal is not fused and so is detected by a general detection model. The following experiment was designed to answer this question.

Experiment 4: Data were collected to compare signal detectability among three stimulus configurations: 1) all components in the masker and signal are harmonically related (harmonic configuration); 2) the masker is a harmonic complex, but the signal is not harmonically related to the masker (nonharmonic configuration); and 3) none of the masker or signal components were harmonically related (arbitrary configuration). Signal detectability in the arbitrary configuration, in which neigther segregation or fusion is likely to occur, serves as a reference to measure the fusion in the harmonic configuarion and the segregation in the nonharmoinic configuration. The signal was a 660 Hz sinusoid. The masker contained either 1, 2, or 3 tones, which were all either harmonic with the signal, harmonic with each other but not the signal, or not harmonic with each other or the signal. The signal and masker were gated together with a 125 ms duration and overall level of 75 dB SPL.

Experimental Series V

The experiments conducted in the 4th and 5th years were an outgrowth of the data collected in the first years of this grant period and represent an attempt to verify that we have, in fact, been studying the acoustic cues used by the auditory system for sound source segregation. We, along with other researchers who are studying sound source segregation, have made reasonable assumptions, based on the acoustics of environmental stimuli (e.g. speech, music), that our experimental paradigms are replicating, to some extent, the process by which the auditory system separates components of individual simultaneous signal into groups. If indeed that is what we are studying, then it follows that individuals who cannot perform the process of sound source segregation will have difficulty with the several experimental paradigms which we think measure this process. With that in mind, last year we took our research in a somewhat different direction.

We have identified two groups of people whose ability to analyze the acoustic environment seems not to include the ability to isolate components of a single sound source from components of competing sounds. The first of these groups consists of those we call "noise-problem people." These are people with normal hearing sensitivity, as determined by pure-tone audiometry, and an absence of ear pathology, but who have enough difficulty listening in noise that they schedule an audiological evaluation. In other words, in spite of hearing that is audiologically normal, they have problems processing the acoustic environment. The second group represents those who are thought to have central auditory processing disorders (CAPD). This is a frequently used term referring to individuals, usually children, who have difficulty listening to a signal in the presence of competing signals. The clinical test battery used to determine the presence of a CAPD is basically a series of tests (monotic and dichotic) to determine if a background noise creates more than the normal amount of difficulty in understanding a speech signal. Our subject pool is made up of people from these two groups.

Our experimental battery consists of several paradigms we had previously used with normal hearing subjects in studying sound source segregation. These were detailed in the last progress report and are based on comodulation masking release (CMR), cross-spectrum fusion, monaural echo suppression, and cross-spectrum frequency change coherence. We will be adding a new paradigm based on simultaneous gating of signal and masker.

We have completed collection of data on 20 normally hearing young adults, ages 18-23. This was necessary since published data on these tasks use subjects who have had varying experiences on these tasks, but who, for the most part, are very experienced. We have followed our experimental paradigm with these subjects and now have normative data on all our experimental tasks, including simultaneous gating. These normative data are listed below:

Comodulation masking release - 12 dB

The signal was a 2000 Hz sinusoid, the masker band, a 100-Hz wide noise band centered at 2000 Hz, and the cue band, a 100 Hz wide noise band centered at 1000 Hz. The cue band was 40 dB higher in level than the masker band. The masker and cue band envelopes were either synchronous or independent, and the difference in detectability between those two conditions is the CMR.

Monaural echo suppression - 20 dB

The signal was a 500-Hz sinusoid which was either leading or delayed by 30 ms relative to the masker, which was also a 500-Hz sinusoid. Detectability of the leading signal is substantially better than that of the delayed signal. When the signal and masker are different in frequency, the difference in detectability between signal leading and delayed virtually disappears in normal hearing listeners.

Frequency change coherence - 8 dB

The two-component masker consisted of two downglides centered at 220 Hz and 440 Hz. The signal was centered at 660 Hz and was either an upglide or a downglide. Data with normal hearing subjects indicate that the signal is easier to detect by 8 dB when it is an upglide.

Cross-spectrum fusion - 16 dB

Signal is an amplitude-modulated 100-Hz wide noise band centered at 2000 Hz. Masker is similar, but centered at 1000 Hz. When the two noise bands are synchronously modulated, detectability is poorer by 16 dB, than when they are independently modulated.

Simultaneous gating - 5 dB

The 50 ms signal was a 1000-Hz sinusoid and the masker was a 500-Hz sinusoid, either 50 ms or a350 ms in duration. They were either gated simultaneously, or the signal was gated on after the masker had been on for 150 ms. Simultaneously gating caused the signal to be 5 dB less detectable.

In addition to the data above, all subjects in the normally-hearing group had audiological evaluations and a clinical auditory test battery (low pass filtered speech, time compressed speech, Staggered Spondaic Word Test and competing sentences). They also completed the questionnaire developed to access difficulty -listening in noise.

We have completed our test battery on each ear of seven subjects who are “noise-problem people” and on 1 subject who has been diagnosed with CAPD.

Preliminary results show that each of our subjects had difficulty on at least one of the experiments. For example, Subject MD (25 years old) had no measurable CMR in his left ear, but a normal CMR in his right ear. His left ear also showed no sensitivity to FM coherence, and minimal sensitivity to echo suppression. His score on cross-spectrum fusion was robust. In his right ear, all data appeared close to normal except for the echo suppression.

Interestingly, CMR seemed to be the most difficult for these subjects. All but one showed a greatly diminished CMR in at least one ear. Four of the subjects showed diminished ability for cross-spectrum fusion, although the other subjects did not. For one (CP) of the subjects, cross-spectrum fusion was his only diminished perceptual skill.

Marion F. Cohen Grant No. AFOSR F49620-93-1-0070

Meeting Papers:

Chen, X. and Cohen, M.F. (1993). Monaural echo suppression: Detectability of a delayed signal based on vowel formant structure," J. Acoust. Soc. Am., 93, 2350.

Manuscripts:

Chen, Xiaofen and Cohen, Marion F. "Acoustic factors influencing detectability of a delayed monoral signal," J. Acoust. Soc. Am. (submitted).

Cohen, Marion F. and Chen, Xiaofen. "Effect of harmonicity on detectability of a masked signal."