Notes and Comment

Comments on Deutsch's musical scale illusion

TEDD JUDD

*Cornell University, Ithaca, New York 14853
and Aphasia Research Center
Boston Veterans Administration Hospital
Boston, Massachusetts 02130*

The fields of auditory perception and music were challenged with a series of intriguing paradoxes when Deutsch (1975a) announced the discovery of a new "musical illusion." The stimulus for this illusion is a dichotic presentation of simultaneous ascending and descending C major scales where each scale switches ears on each change of note as follows:

Channel 1 . . . 1 7 3 5 5 3 7 1 1 7 3 . . .
Channel 2 . . . 8 2 6 4 4 6 2 8 8 2 6 . . .

(1 represents the lowest and 8 the highest notes of the scale.) The tone durations were 250 msec. The results were complex, but Deutsch's major findings were that most subjects reported hearing the following two streams or melodies:

Stream 1 . . . 8 7 6 5 5 6 7 8 8 7 6 . . .
Stream 2 . . . 1 2 3 4 4 3 2 1 1 2 3 . . .

Furthermore, the majority of the right-handed subjects reported that they heard Stream 1 in the right ear and Stream 2 in the left ear regardless of which channel was presented to which ear.

I am presenting here an explanation of the tonal pattern organization aspects of Deutsch's illusion based on recently developed rules of auditory pattern perception and known principles of auditory localization.¹ This explanation is supported by the perceptual reports resulting from several manipulations of Deutsch's stimuli. I will not offer any explanation of the lateral asymmetries Deutsch reported.

The perceptual rules on which this explanation is based are those of auditory stream segregation. When rapid tone sequences (about 1-20 tones/sec.) are presented in which successive tones vary widely along such dimensions as pitch (Bregman & Campbell, 1971),

location of sound source (Judd, Note 1), loudness (van Noorden, 1975), and possibly timbre, attack, and rhythm (Judd, Note 2), listeners tend to perceive the tones as grouped along these dimensions, i.e., they may hear a high-pitched melody approximately simultaneously with a low-pitched melody, a melody on the left approximately simultaneously with a melody on the right, or a loud melody approximately simultaneously with a soft melody. This separation of a single sequence of notes into two or more perceived melodies or streams is called stream segregation (SS) (van Noorden's "fission"). The term is modified according to the dimension of variation, i.e., pitch stream segregation (PSS), spatial stream segregation (SSS), loudness stream segregation (LSS), etc. When the tone presentation rate is decreased and/or the separation along the dimension of variation is decreased, the sequence is heard as a unified whole, i.e., there is pattern integration (van Noorden's "temporal coherence"). At intermediate speeds and separations, the stimulus can generally be perceived in either mode depending on the attentional bias of the listener and possibly other factors. This range of stimulus parameters is designated the ambiguous region. These relationships are summarized in Figure 1. (Note that the perceptual phenomena are actually more complicated for loudness; see van Noorden, 1975.)

Deutsch's illusion involves a stimulus in which the information favorable to a perceptual organization based on pitch range is in conflict with information favorable to a perceptual organization based on the spatial separation of the sound sources. In other words, if the notes are heard as organized by pitch, a particular pair of melodies or streams should be

Figure 1. Summary of auditory stream segregation phenomena. Adapted from van Noorden (1975) and Judd (Note 1, Note 2). Values for the "separation of tones" (in decibels, semitones, angles of separation in space relative to listener, etc.) cannot be applied because of limitations in the ranges of available data and method differences between the studies.

This research was supported by N.I.M.H. Training Grant MH-13894-03 to the Cornell University Department of Psychology. I am grateful to James Farber and William Hemsath for technical assistance, to Steffanie Shattuck-Hufnagel, Marlene Oscar-Berman, and Nancy Vanderveer for comments on earlier drafts, and to Claire Sybertz for help with the manuscript. This paper is dedicated to the memory of Eric Lenneberg.

heard, while if they are heard as organized by space, a different pair of melodies should be heard. There is a conflict because each of the space-organized melodies consists of notes from both of the pitch-organized melodies, and vice versa. We can predict on the basis of SS phenomena that the frequency with which each of these experiences occurs should reflect not only the relative strength of each of these organizing tendencies, but also the salience of the information supporting each tendency.

Deutsch's stimulus does not favor perceptual organization by space because it is difficult to localize the signal at each ear separately. There are three sources of information for the lateralization of sound in headphones presentation: interaural sound arrival time differences, phase differences, and interaural intensity differences (Mills, 1972). With Deutsch's stimulus, each ear receives only one of the two signals and the signals arrive simultaneously, therefore neither signal can be lateralized on the basis of interaural arrival time differences or phase differences. What little data is available concerning intensity-based lateralization of dichotic tones comes from studies which were directed at different questions (Efron & Yund, 1974; House & Wohlford, 1975), or had a narrow range of pitch separation (e.g., van den Brink, Sintnicolaas, & van Stam, 1976), or were performed using much higher pitches (Scharf, Note 3) where localization and pitch perception both show different properties from the range where the Deutsch study was performed. Nevertheless, these studies suggest the tentative conclusion that simultaneous dichotic tones close to one another in pitch are likely to be perceived as coming from a single location, whereas tones of very different pitch are independently and accurately localized. The range of pitch separations which result in a single localization extends well beyond the range of pitch fusion (including binarual beats, rotating tones, etc.) where a single pitch is perceived. The single lateralization is towards the tone with the greater subjective intensity. There is some indication (Deutsch, 1974; House & Wohlford, 1975; and the author's unpublished observations) that a single localization may be more common for tones an octave apart than for tones somewhat closer together (compare to Scharf, Note 3). Considering the close range of dichotic pitches in Deutsch's (1975a) scale illusion, their equal objective intensities, and the possible special qualities of octaves in this regard it seems likely that all lateralization will be difficult and unreliable and that independent lateralization of the tones at each ear will be especially difficult. The low salience of the information for separate tone localization will make perceptual organization by space (SSS) very unlikely.

On the other hand, there is little evidence to suggest that there is significant interference with the salience of pitch perception with Deutsch's peculiar stimulus

pattern. Binaural diplacusis—the discrepancy of perceived pitch across the two ears—generally stays within the range of a musical half-step for the general population with normal hearing (Ward, 1970). Yund and Efron (1975) found that the suppression of pitch information from one ear in a dichotic presentation of two-tone sequences was present predominantly at the extremes of intensity differences between the ears (50 dB). Other investigators, using related paradigms, apparently failed to inquire about the presence of both pitches (Deutsch, 1975b; House & Wohlford, 1975; Licklider, 1967). Hence, we can expect that perceptual organization by pitch (PSS) will be very likely to occur.

Two tests were generated in support of the above line of argument as an explanation of Deutsch's scale illusion. The first tests the idea that dichotic presentation of sounds in the same frequency range decreases the likelihood of SSS through a reduction in the salience of lateralization information. In the second test Deutsch's stimulus pattern is altered to restore lateralization information. A consequent reorganization of the perception of the pattern based on spatial information (SSS) was predicted.

STUDY 1

In an earlier study of SSS (Judd, Note 1), I found the following: (1) SSS is a function of rate of tone presentation and the separation of the tones in perceptual space, as summarized in Figure 1. (2) In a tonal pattern identification task, it was difficult to identify sequences which were distinguished only by the temporal order of tones between two different spatially segregated streams.

In addition, I found that when subjects were experiencing SSS while listening to a stimulus which could be perceived either as pattern integration or as SSS (the "ambiguous region" of Figure 1), pattern integration could be induced by several events which introduced excess noise, such as yawning, talking, or breathing deeply, as well as by events which reduced the salience of localization, such as moving the head (with a stimulus presentation via speakers) or day-dreaming. The reverse was not the case, although subjects frequently went from pattern integration to SSS by concentrating on the localization of the sounds.

I proposed the following explanation for this phenomenon: SSS is dependent upon the spatial localization of the stimuli, which, in the case of stimuli alternating between the ears via headphones, is determined entirely by the loudness difference between the two ears. Any change in the stimulus situation which reduces the salience of auditory localization ought to reduce the probability that ambiguous stimuli will be perceived in a SSS fashion. The introduction of binaural noise described above ought to have such an effect if, as the results of Efron and Yund (1974)

and Scharf (Note 3) suggest, localization of any one sound involves a loudness comparison between the ears over a range of frequencies.

To test this explanation, the following experiment was devised: performance on the original tonal pattern identification task was compared to performance on the same task when the silence in the ear opposite each tone was replaced by noise with an amplitude equal to that of the tone. I predicted that this would decrease localization of the sounds, decrease SSS, and improve performance on discrimination of the two patterns.

Methods

Subjects. The subjects were all unpaid adult volunteers with normal hearing by their own reports. Some subjects had participated in another auditory pattern-perception study (Judd, Note 1) but were unaware of the purpose of this study.

Stimuli and Apparatus. The task was a same-different discrimination similar to the second study by Judd (Note 1). Two stimulus patterns were constructed out of the following four triangle-wave tones: (1) 912 Hz (musical A# on the well-tempered scale), (2) 1,024 Hz (C), (3) 1,150 Hz (D), (4) 1,290 Hz (E). (These tones are close enough to one another in pitch so as to make stream segregation by pitch unlikely; van Noorden, 1975.) Each tone was 100 msec long with 5-msec exponential onset and offset amplitude ramps to minimize clicks. The only overlap between successive tones was on these ramps. The patterns were as follows ("S" indicates silence).

Pattern A Channel 1 . . . 1 S 4 S 1 S 4 . . .
Channel 2 . . . S 2 S 3 S 2 S . . .

Pattern B Channel 1 . . . 1 S 4 S 1 S 4 . . .
Channel 2 . . . S 3 S 2 S 3 S . . .

Note that the tone sequence on each channel taken separately is the same for both patterns (1 S 4 S and 2 S 3 S); only if a temporal comparison between the two channels is made can the two patterns be distinguished. In the noise condition, the silences in the channel opposite to where a tone was being presented (all of the Ss in the diagrams above) were replaced by 20-Hz-20-kHz white noise with an RMS signal amplitude equal to that of the tone.

Each subject was run for one session. The session consisted of eight trial blocks of five trials each. Noise and no-noise conditions were on alternate blocks, beginning with the noise condition, so there were 20 trials in each condition. The interblock interval was 8 sec, the intertrial interval was 2 sec, and the interstimulus interval was 500 msec. Each stimulus consisted of 300 msec each of onset and offset amplitude ramps and 2 sec of full amplitude stimulus. Each stimulus began at a random point in the cycle.

The stimuli were recorded on tape using a Revox A77 tape recorder. The noise was produced by a Lafayette 15011 white-noise generator. The tones were generated using the voltage-controlled oscillator of a Minimoog synthesizer. This was driven by a Moog 960 sequential controller, which also drove the MMC-4A voltage-controlled amplifier switches assigning the tones and noise to different channels of the recorder. Frequencies, durations, and amplitudes were checked periodically as the tape was made using a Hewlett-Packard 5302A Universal counter and a Ballantine 323-01 RMS voltmeter. The stimuli were delivered over Sharpe 660 headphones in an acoustic isolation chamber.

Procedure. The task and stimuli were fully explained to each subject individually before the experiment, including all of the

information given above and the diagrams but excluding numerical values. The one piece of information excluded from the stimulus description was the exact location of the noise within the stimulus patterns; the subject was simply told, "There will be some noise added to the patterns in the odd number trial blocks." Overall loudness was adjusted to the subject's comfort from a random point on the tape before the experimental session began; this always fell between 50 and 80 dB SPL for the full-amplitude stimulus. The tape began with an introduction in which each pattern was played, beginning at a slow speed with tone durations of 1,000 msec and slowly speeding up to 100-msec tone durations. The subject responded by writing "S" or "D" for "same" and "different" on an answer sheet according to whether they thought the pairs of tone sequences they heard had the same or different orders of pitches.

Subjects who had done very well in Judd's Study No. 2 (J.M.; N.V. J.J.) and self-professed musicians were presented with the tape at the speed of recording. All others (designated by "*" under "Half Speed" in Table 1) were given the tape at one-half the recording speed, thus doubling all of the time values given above and halving the frequencies (dropping them an octave). This was done because the Judd studies had shown large individual differences in the ability to hear tonal patterns such as this as an integrated whole and to perform tasks similar to this.

After the experiment, each subject was asked for general comments about the experience and, in particular, was asked to compare the two experimental conditions. Finally, the subject was asked to compare the two conditions on the ease with which the tones were heard as a single, integrated series.

Results

The results are presented in Table 1. Eight of the 12 subjects performed better on the noise than on the no-noise condition, as predicted by the hypothesis. This difference was significant at the $p < .05$ level, two-tailed by the chi-square test, for 3 subjects and at the $p < .01$ level for all 12 subjects taken as a group. Four subjects performed better on the no-noise condition than on the noise condition, but this difference was not significant for any of them. These

Table 1
Performance on Across-Ears Temporal Order
Judgment Task in Study 1

Subject	Half-Speed	Phenomenology	Percent Correct (20 Trials per Condition)		
			Noise	No Noise	Difference
F.F.**	*	*	100	50	50†
J.M.		*	100	60	40†
J.K.	*	*	95	75	20†
D.C.	*	*	65	45	20
T.J.		*	70	60	10
N.V.			90	80	10
S.S.	*	*	75	70	5
K.L.	*	*	60	55	5
J.P.	*		75	80	-5
M.H.			95	100	-5
D.K.	*		70	80	-10
J.J.			75	90	-15
Sum	7	7	80	71	9††

*See text for explanation.

**Ten trials per condition.

† $p < .05$, two-tailed test.

†† $p < .01$, two-tailed test.

subjects were among those who afterwards reported no difference between conditions in their ability to hear the patterns as an integrated whole; these subjects typically reported no difficulty in either condition in hearing the patterns as an integrated whole.

Seven of the subjects (designated with a "*" under "Phenomenology" in Table 1) reported that it was "easier" or "much easier" to hear the patterns as an integrated whole during the noise condition than during the no-noise condition. All of these subjects gave this report before being asked to make that particular comparison, which is a testimony to the salience of the phenomenon for these subjects. No subject reported the opposite, even after the final question.

Discussion

The results of this experiment are consistent with the hypothesis that the presentation of noise to the ear opposite that receiving the tone increases the likelihood of pattern integration and improves performance on a discrimination based on across-ears temporal ordering. The proposed explanation predicting such a result, based on the degradation of localization information was supported by the failure of many subjects to properly localize the noise. (A similar report regarding the tones could not be obtained, since the subjects were told in advance where the tones would be.)

The great individual differences in the speeds at which SSS occurs may help to explain why some subjects failed to report easier pattern integration with the noise and showed insignificant differences in performance between the conditions. The performances and reports of these subjects suggest that they had little difficulty hearing any of the patterns as integrated.

This experiment and its results are very similar to Schubert and Parker's (1955) experiment with speech alternated between the ears. In the original experiment with alternating speech, Cherry and Taylor (1954) found that performance on a shadowing task was worst at a switching rate of 3 Hz, with better performance at slower and faster switching rates. Schubert and Parker replicated this effect and compared it to performance on the same task when the stimulus included noise played to the ear opposite the speech at about the same amplitude, similar to the noise in the ear opposite to the tone in this study. They were amazed to find that the noise virtually eliminated the rate-dependent effect and, in fact, improved performance at all rates. They could think of no explanation for this result, but speculated that the noise may have eliminated some stimulus offset inhibition. The present experiment suggests, instead, that the noise degraded the information for the localization of the speech and decreased SSS. [Egan (1948) found that adding noise to the ear opposite monotonically presented

speech resulted in an impression of either less precise localization of the speech or movement of the speech towards the center of the head.] The better integrated perception of the whole speech signal resulted in an improved performance on the shadowing task.²

STUDY 2

I have demonstrated that auditory localization and SSS can be hampered by certain conditions of dichotic stimulation. In this study, I demonstrate that SSS can be restored to the perception of stimuli similar to Deutsch's (1975a) scale illusion at the expense of PSS if the stimuli are restructured to give a nondichotic presentation, allowing accurate lateralization by interaural loudness differences. This restructuring was achieved by having the notes alternate between the ears in two different ways while keeping the overall cycle time of the pattern constant.

Method

Subjects. Some of the 24 subjects were the same as in Study 1; all were adults with normal hearing by their own reports and were unaware of the purpose of the study. Fifteen were right-handed and nine reported left or mixed handedness.

Stimuli and Apparatus. The equipment was the same as above. For practical reasons, the Deutsch stimulus was truncated to one-half its original length, triangle waves were used instead of sine waves, and 10-msec overlapping exponential amplitude ramps were added to both ends of the notes in Patterns 2 and 3 to avoid transients; otherwise all parameters were identical to Deutsch's. The following three patterns were generated where "S" indicates silence and the digits represent these triangle-wave tones: (1) 259 Hz (musical C), (2) 290 Hz (D), (3) 326 Hz (E), (4) 345 Hz (F), (5) 388 Hz (G), (6) 435 Hz (A), (7) 488 Hz (B), (8) 517 Hz (C):

Pattern 1 Channel 1 . . . 1 7 3 5 1 7 3 5 1 . . .
Channel 2 . . . 8 2 6 4 8 2 6 4 8 . . .

Pattern 2 Channel 1 . . . 1 S 7 S 3 S 5 S 1 . . .
Channel 2 . . . S 8 S 2 S 6 S 4 S . . .

Pattern 3 Channel 1 . . . 1 S S 7 3 S S 5 1 . . .
Channel 2 . . . S 8 2 S S 6 4 S S . . .

The tones in Pattern 1 were 250 msec long, while the tones in Patterns 2 and 3 were 125 msec long to maintain an overall cycle time of 1 sec as in Pattern 1.

In both Pattern 2 and Pattern 3, the tones in each channel occur in the same order as they occur in that channel in Pattern 1. Furthermore, each note occurs immediately before or after the note it was simultaneous with in Pattern 1; in other words, one half of each note from Pattern 1 has been eliminated, while the remainder retains its absolute position. In Pattern 2, there is a strict alternation of notes between the channels, while in Pattern 3, two notes occur in succession in one channel followed by two notes in the other channel, which allows for a strict alternation of the ascending (1234) and descending (8765) components.

Procedure. Forty-five seconds of each pattern was recorded on tape. The subjects were presented with the patterns on headphones. Overall loudness was adjusted to each subject's comfort (always

Table 2
Response Frequencies of 24 Subjects to Variants of Deutsch's (1975a) Scale Illusion in Study 2

	Pattern 1		Pattern 2		Pattern 3	
	Stereo	Binaural	Stereo	Binaural	Stereo	Binaural
Organization by Ear						
Incomplete	9(6)	0	4(1)	2	16(3)	0
Complete	1	0	19	0	6(1)	0
Organization by Pitch						
Incomplete	11(3)	12	3(2)	0	7(6)	2
Complete	12(1)	11	0	22	4(1)	21
Organization by Whole	0	0	1	11	0	2
Illusory Localization	17	4	1	4	4	2
Correct Localization*		4		2		7

Note—For stimulus patterns, categorization criteria, and further explanations, see text.

*Correct localization with organization by pitch or whole.

between 50 and 80 dB SPL). Each pattern was presented in a stereo condition and in a binaural condition with both channels combined. Channels were assigned to ears randomly and noted. Order of condition presentation and order of pattern presentation were counter-balanced across subjects. Both conditions for each pattern were presented before proceeding to the next pattern.

The subjects were told that they would hear repeated patterns of notes, and they were asked to describe orally what they heard as best they could. Singing was allowed. If the end of the tape segment was reached before they had finished their description, it was repeated. The experimenter faced the subject and took notes on his/her descriptions, monitoring the tape by the faint leak from the headphones.

If the subject's complete description was inadequate to determine what information s/he had used in grouping the tones s/he was asked to describe what patterns s/he had heard. If after that s/he still had not spontaneously reported a localization for the patterns s/he heard, s/he was asked to indicate where each pattern was heard; care was taken to avoid suggesting that a lateralization would be considered an appropriate response. The trial was ended when all of the questions had been answered to the best of the subject's ability and the subject was satisfied with his/her description of the pattern.

Results

The following criteria for the categorization of the responses were established post hoc: *Organization by ear*: Incomplete—The subject reported at least three tones from one ear or two tones from each ear with correct pitch contour and rhythmic relations within each ear and correct localization. Complete—The subject reported all eight tones ear by ear with the correct pitch contour and rhythmic relations within each ear and correct localization. *Organization by pitch*: Incomplete—The subject reported at least three tones or two groups of two tones close together in pitch, drawing from both ears (in the case of stereo presentation), with correct rhythmic and pitch contour relations within each group. Correct localization was not required. Complete—The subject correctly accounted for all notes in groups classed by pitch where at least one group included notes from both ears (for the stereo conditions) with correct rhythmic and pitch contour relations. Correct localization was not re-

quired. *Organization by whole*: All eight tones were reported as a single group.

The frequencies of each category of response for each condition are summarized in Table 2. The subjects sometimes said they could hear a stimulus in two different ways. When a response fell into two categories, both were added into the totals; the portion of the total contributed by the second part of the subjects' reports is presented in parentheses in the affected cells.

All subjects but one heard an organization by pitch for the stereo version of Pattern 1, although it was incomplete for 11 of them. This usually meant that they reported only one of the two patterns of four notes (Deutsch's "single-stream" subjects). Unlike Deutsch's single-stream subjects, however, 4 of these subjects reported only the lower ascending sequence. Sixteen subjects showed some tendency to lateralize these pitch patterns. Excluding 3 subjects who lateralized both patterns in one ear and/or one pattern in both ears, these results are reported twice in Table 3 once by ear-to-channel assignments and once by stream-to-ear lateralizations. Ten subjects also showed

Table 3
Illusory Lateralization of Pattern 1, Stereo Version, in Study 2

Lateralization*	Subject Handedness	
	Left or Mixed	Right
Channel 1	4	4
Channel 2	2	3
Right Ear	2	3
Left Ear	4	4

Note—Data reflect the number of subjects (out of 24) reporting an illusory lateralization of a given portion of Pattern 1 toward a given channel and toward a given ear by subject handedness. For stimulus patterns and channels, see text. *Ascending pattern lateralized toward one channel or ear (and/or descending pattern lateralized toward the opposite).

some tendency to hear patterns organized by ear in this condition, although these were incomplete in all but one case. One subject gave a veridical report of all pitches and their localizations. With the binaural version, virtually all subjects heard organizations by pitch.

In the stereo condition, an organization by ear was reported by all subjects but 1 for Pattern 2 and by all subjects but 2 for Pattern 3, although this was less complete for Pattern 3. Some organization by pitch was reported by 3 subjects for Pattern 2 and by 11 subjects for Pattern 3. Eight of the latter reported the correct localization for these patterns, i.e., alternating ears. Almost all subjects readily heard the binaural version of Pattern 3 as an ascending and a descending group, each of four notes, with the descending group higher than the ascending group. The organization of the binaural version of Pattern 2 was less pronounced and less consistent. The most common report was of a three-note ascending low group (1_23_) and a five-note descending high (_87_ 654). Four subjects showed a fleeting tendency to lateralize a pattern organized by pitch in the stereo version for Pattern 3. In all cases, this consisted of the lateral displacement towards Channel 1 of the middle note in a three-note pattern. Six subjects reported a lateralization of pitch patterns in one or more of the binaural conditions.

Discussion

Altering Deutsch's stimulus so as to restore localization information resulted in the dominance of the SSS mode of perceiving the stimulus over the previously dominant PSS mode for all subjects. Thus the tonal patterns reported by Deutsch's subjects are simply those patterns which are predicted by stream segregation theory and the principles of auditory localization.

There are at least two reasons why Pattern 2 shows a tendency towards SSS and away from PSS compared to Pattern 3. First, in Pattern 2 the spatial switching rate is twice that in Pattern 3, while roughly the opposite may be said for the pitch range switching rate. There is a greater tendency towards SS at higher switching rates (Figure 1; Judd, Note 1; van Noorden, 1975). Second, in Pattern 2 the spatial streams show a stronger rhythmic organization than the pitch streams which may make the spatial streams more salient, while in Pattern 3 the opposite is the case.

The questions that I asked my subjects regarding localization of the tones and tonal patterns were open-ended, and I found a much greater variability in the types of responses given than were found by Efron and Yund (1974) with their forced choice paradigm, by Deutsch (1974, 1975a, 1975b) with her suggested ear-by-ear report, or by House and Wohlford (1975) with their directed questioning. For example, five of my subjects heard the ascending pattern from the

stereo condition of Pattern 1 in the center of the head and six subjects reported a lateralization of binaural stimuli. Whether these differences in response type variability can be attributed to procedural or stimulus differences between these studies cannot be determined from the available data.

My findings on laterality do not resemble any of the conflicting laterality results of the studies mentioned above. A comparison of these results would be pointless in view of the differences in procedure. One conclusion can be safely drawn, however: more work is needed in this area utilizing large numbers of subjects, a variety of verbal and tonal materials, and careful, open-ended questions which will not cover up the real variety of perceptions found in different subjects.

SUMMARY AND CONCLUSIONS

Deutsch has presented us with the first experimental example of a conflict between PSS and SSS; the full range of interactions between these phenomena remains to be explored. Nevertheless, the reasons for Deutsch's result—the domination of PSS over SSS—are already apparent. Three stimulus configuration characteristics are responsible: (1) the information favoring PSS is in conflict with the information favoring SSS; (2) the information favoring SSS is poor; and (3) the information favoring PSS is good.

(1) The conflict between the information for PSS and that for SSS is clear because the ranges of the pitches presented to each ear overlap. I have found (unpublished observations) that the confusion between the ears can be eliminated by doubling the frequencies of the tones presented to one ear in Deutsch's stimulus (thereby raising the pitch of these notes one octave and removing them from the range of the notes in the other ear).

(2) The information favoring SSS is auditory localization information. This has been discussed in detail above.

(3) Deutsch's upper and lower streams obey the centuries-old law of counterpoint specifying that melodic lines should not cross (Piston, 1947), which Dowling (1973) has found to be sufficient to allow the perceptual segregation of two different melodies or streams.

As Deutsch (1975a) has noted, good form is also a factor which probably favors PSS in the scale illusion. When the collection of notes comprising my Pattern 1 is split in half by pitch range, the result is a simple ascending pattern of four notes and a simple descending pattern of four notes. In contrast, when they are split by ear of presentation, the result is an up-down-up-down pattern. Wilcox (Note 4) has shown that patterns with one direction of frequency change are simpler than sequences which show shifts in the direction of frequency change, in that

it is easier to judge the order of tones in repeated sequences of the first type. I have found (unpublished observations) that making the tone patterns in each channel less "difficult" by these criteria is sufficient to reinstate SSS. If the Deutsch stimuli are not switched from ear to ear, that is, if simultaneous ascending and descending scales are presented dichotically or on widely separated speakers like this:

Channel 1 . . . 1 2 3 4 5 6 7 8 1 . . .
Channel 2 . . . 8 7 6 5 4 3 2 1 8 . . .

it is easy to hear each scale segregated and in its entirety. However, if the speakers are moved closer together, there comes a point where PSS becomes more salient. This occurs well before it becomes impossible to independently localize the two speakers (90° for me) so that each of the pitch-segregated patterns seems to move back and forth between the speakers once in each cycle.

Deutsch (1975a) was surprised that pitch information should dominate over spatial information in her study because much research on listening to two or more simultaneous speech messages has indicated that, "The most effective dimension so far discovered along which a message may be located for selection is without doubt position in auditory space" (Moray, 1969). This seems less surprising in light of the present findings, because only very special stimulus conditions have produced this dominance of PSS over SSS. In most dichotic speech studies, particularly those in the attention literature from which Moray drew his conclusions, there is much more information for localization than in the Deutsch illusion. Each channel contains a signal with a complex, unique, and rapidly changing amplitude envelope which affords many opportunities for detecting loudness differences between the ears. Furthermore, Moray was referring to a large number of studies using speakers rather than headphones so that phase and onset-time differences between the ears are available as additional information for localization.

In addition, many of those studies used running speech, so the structure and redundancy of the messages probably contributed to their perceptual separation. In fact, if the message contexts are similar in the two channels of a dichotic speech stimulus, fusion between the two channels can occur, e.g., "PAY" and "LAY" delivered to opposite ears is often perceived as "PLAY" (Cutting, 1975).

Several of the subjects in Study 2 found listening to the stimuli very enjoyable. They found that with repeated listening they could direct their attention to different ways of organizing the tones, especially with the stereo version of Pattern 3. These attentional phenomena are worthy of investigation in their own rights. Many of the classic laws of musical counter-

point (Piston, 1947) seem to place the different contrapuntal voices into the "ambiguous region" of Figure 1 with respect to one another on a variety of dimensions. In other words, contrapuntal music is frequently composed so as to allow the listener to choose between attending to single lines or to the relations between lines. The many listening alternatives contribute to the intriguing quality of the music. This also suggests that the stimulus techniques and organizational principles discussed here could prove useful in the composition of electronic music for headphone listening.

REFERENCE NOTES

1. Judd, T. *Auditory spatial stream segregation*. Paper presented at the 97th meeting of the Acoustical Society of America, Cambridge, Massachusetts, 1979.
2. Judd, T. *Some new forms of auditory stream segregation and their interactions*. Manuscript in preparation.
3. Scharf, B. *Frequency selectivity and sound localization*. Paper presented at the Symposium on Hearing Theory, Eindhoven, Holland, 1972.
4. Wilcox, G. W. *Temporal coherence of tone sequences*. Paper delivered at the American Psychological Association meeting, Honolulu, 1972.

REFERENCES

- BREGMAN, A. S., & CAMPBELL, J. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 1971, **89**, 244-249.
- VAN DEN BRINK, G., SINTNICOLAAS, K., & VAN STAM, W. S. Dichotic pitch fusion. *Journal of the Acoustical Society of America*, 1976, **59**, 1471-1476.
- CHERRY, E. C., & TAYLOR, W. K. Some further experiments upon the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 1954, **26**, 554-559.
- CUTTING, J. E. Aspects of phonological fusion. *Journal of Experimental Psychology*, 1975, **104**, 105-120.
- DEUTSCH, D. An auditory illusion. *Nature*, 1974, **251**, 307-309.
- DEUTSCH, D. Two-channel listening to musical scales. *Journal of the Acoustical Society of America*, 1975, **57**, 1156-1160. (a)
- DEUTSCH, D. Musical illusions. *Scientific American*, 1975, **233**, 92-104. (b)
- DOWLING, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, **5**, 322-337.
- EFRON, R., & YUND, E. W. Dichotic competition of simultaneous tone bursts of different frequency—I. Dissociation of pitch from lateralization and loudness. *Neuropsychologia*, 1974, **12**, 249-256.
- EGAN, J. P. The effect of noise in one ear upon the loudness of speech in the other ear. *Journal of the Acoustical Society of America*, 1948, **20**, 58-62.
- HOUSE, A. S., & WOHLFORD, R. E. Some auditory illusions reheard: Dichotic listening to discrete frequencies. *Journal of the Acoustical Society of America*, 1975, **58**, 554 (A).
- HUGGINS, A. W. F. Distortion of the temporal pattern of speech: Interruption and alternation. *Journal of the Acoustical Society of America*, 1964, **36**, 1055-1064.
- LICKLIDER, J. C. R. Phenomena of localization. In A. B. Graham (Ed.), *Sensorineural hearing processes and disorders*. Boston: Little, Brown, 1967.
- MILLS, A. W. Auditory localization. In J. V. Tobias (Ed.), *Foundations of modern auditory theory II*. London: Academic Press, 1972.

- MORAY, N. *Attention*. London: Hutchinson Educational, 1969.
- PISTON, W. *Counterpoint*. New York: Norton, 1947.
- SCHUBERT, E. D., & PARKER, C. D. Addition to Cherry's findings on switching speech between the two ears. *Journal of the Acoustical Society of America*, 1955, **27**, 792-793.
- VAN NOORDEN, L. P. A. S. *Temporal coherence in the perception of tone sequences*. Unpublished doctoral dissertation. Institute for Perception Research, Eindhoven, Netherlands, 1975.
- WARD, W. D. Musical perception. In J. V. Tobias (Ed.), *Foundations of modern auditory theory I*. London: Academic Press, 1970.
- YUND, E. W., & EFRON, R. Dichotic competition of simultaneous tone bursts of different frequency—II. Suppression and ear dominance functions. *Neuropsychologia*, 1975, **13**, 137-150.

NOTES

1. Deutsch (1975a) discussed the illusion as an example of stream segregation in her initial report and noted that it represented a domination of channeling by frequency over channeling by spatial location. However, she gave no reasons for this domination, nor did she give any consideration as to how the situation might be reversed. The contribution of this study, then, is the

application of stream-segregation principles to explain this domination and predicts its reversal.

Dr. Deutsch informs me that she has been conducting similar experiments independently in pursuit of a similar explanation. See Deutsch, D. Binaural integration of tonal patterns. *Journal of the Acoustical Society of America*, 1978, **64**, S146.

2. This explanation implies that the original Cherry effect is due to SSS. On the other hand, Huggins (1964) has shown that the Cherry effect is a function of speech rate and therefore any explanation based simply on an attention-switching hypothesis is inadequate. I agree that a SSS explanation will not account for Huggins' speech rate-dependent effect. However, it is necessary to explain why there should be any disturbance in speech perception at all when speech is alternated between the ears, that is, there must be some reason why the speech segments cannot be centrally reabuted to produce an accurate and complete perception. I maintain that SSS furnishes such an explanation. The rate of alternation at which unavoidable SSS begins is 2-3 Hz for speakers with 150° separation (Judd, Note 1), and this is the rate at which shadowing performance with alternated speech begins to degrade.

(Received for publication May 9, 1978;
revision accepted February 11, 1979.)