Congruency Effects with Dynamic Auditory Stimuli: Design Implications

Bruce N. Walker and **Addie Ehrenstein**
Psychology Department
Rice University
6100 Main Street
Houston, TX 77005-1892 USA
+1 (713) 527-8101
walkerb@rice.edu, addie@rice.edu

ABSTRACT
Since pitch is a commonly varied parameter in auditory displays, we investigated whether it is possible to attend to relative pitch while ignoring changes in pitch, and whether changes in pitch could be assessed independently of the overall pitch of a dynamic auditory stimulus. Stimuli were defined as either *congruent* (e.g., high pitch stimulus that became higher in pitch) or *incongruent* (e.g., high pitch stimulus that became lower in pitch). In this experiment, faster responses to congruent stimuli indicated a failure of selective attention. This effect was uniform for pitch judgments with all stimuli, but varied with the overall pitch for pitch-*change* judgments. The performance difference between congruent and incongruent trials was greatest for the extreme (high or low) stimuli. Moreover, pitch information intruded more into responses to pitch change than vice versa. Auditory display designers can use congruent stimuli to help distinguish between high and low pitches. If pitch change is the important dimension, designers should restrict the range over which stimulus pitches vary.

Keywords
Congruency, pitch separation, dynamic auditory stimuli

INTRODUCTION
In order to design auditory interfaces which afford better comprehension and elicit faster and more accurate reactions, one must first understand how different attributes of auditory stimuli interact to influence perception and responding. One major question is whether it is possible to attend selectively to a given dimension of a sound, while ignoring other dimensions of the sound. For example, a Geiger counter operator may need to listen specifically to the temporal pattern of the sound, which indicates the prevalence of radioactive particles, and ignore the pitch of the sound, which may indicate the types of particles that are present. On the other hand, in monitoring a landing approach an air traffic controller may need to attend to the rate of pitch change that represents rate of descent of an airplane, while paying less attention to the absolute pitch of the sound, which represents the actual altitude of the plane.

Of the various attributes that may be used to display information, pitch is of primary interest because it is the dimension most commonly used to represent data in auditory displays. In particular, it is important to know how pitch interacts with dynamic changes of the stimulus (including changes in pitch itself) because the auditory representation of data rarely involves single, static values (i.e., unchanging pitches). The present research explores the effects of pitch change on pitch classification, as well as the effects of relative pitch on the classification of pitch change. The results presented here are a subset taken from a larger series of experiments that also investigated stimulus-response relations with dynamic auditory stimuli [18].

Pitch perception
The perception of pitch has been studied extensively, resulting in a wealth of research exploring the physics of sound and the mechanisms of hearing (e.g., [12]), the psychophysical aspects of perception and discrimination of different pitches (e.g., [1], [12], [17]), how pitch fits into the structure of music (e.g., [14]) and the psychological aspects of hearing, and pitch and music perception (e.g., [2]). Pitch is often used as a dimension in auditory displays for the very reason that so much is already known about simple pitch perception. In addition, most listeners are familiar with the concept of pitch and can detect fairly small pitch changes with little training (e.g., less than 3-Hz change in a pure tone of 1000 Hz; [12]). Another reason for using pitch as a display dimension is the relative ease with which pitch can be controlled by current

display hardware [5]. Pitch also more evenly represents a wider range of values than, say, loudness, since at the extremes loudness does not provide an effective display dimension (soft sounds are masked by ambient noise, and loud sounds are potentially disturbing or even damaging).

Pitch in interaction with other dimensions

Relatively little is known about how pitch interacts with other stimulus dimensions and how successful listeners can be at attending to just one of several auditory display dimensions. For many pairs of stimulus dimensions subjects are able to respond solely on the basis of the information contained in the relevant or "cue" dimension, and successfully ignore the other, "irrelevant" dimension. However, for other stimulus dimension pairs subjects are not able to attend selectively to just one dimension. Rather, performance is disrupted by variations in the other dimension. An example of how pitch can affect perception of another stimulus dimension is Melara and Marks' [8] finding that listeners responded faster to a loud sound if the sound was also high, rather than low in pitch—despite instructions to ignore the pitch. Correspondingly, responses to soft sounds were faster if the pitch was low, rather than high. Other studies have demonstrated interactions of pitch with other auditory dimensions, including loudness (e.g., [4], [7], [9], [16]), timbre (e.g., [8]), waveform and duration (e.g., [19]) and the physical location of the sound (e.g., [15]).

Dynamic stimuli and congruency effects

To date, virtually all of the research on interactions of pitch with other stimulus dimensions has involved an unchanging pitch and another static stimulus dimension. Thus, there is still much to be learned about dynamic auditory stimuli, and, in particular, about how changing pitch plays a role in perceiving and making responses to other aspects of a sound. Recent findings with dynamic visual stimuli suggest that changes within a stimulus dimension may interact with responding to a given value of that same stimulus dimension. For instance, it has been shown that perception of the physical position of a visual target is influenced by the direction of motion of the target, even when the task is to respond to, for example, onset position and ignore the motion of the target (e.g., [3], [10], [11], [13]). In the case where the task is to attend selectively to the onset position of a visual target (i.e., whether a square appears on the left or the right side of the display), responses are typically faster if the position (e.g., left) is congruent with the direction of motion of the same visual target (e.g., the square moves farther to the left). Responses are slower if the position and direction of movement are incongruent (i.e., the square appears on the left, but moves toward the right side of the display; [3], [13]). It remains to be seen whether, in the auditory domain, the interaction of pitch and pitch change produce similar congruency effects, so that if a sound is high in pitch, responding is faster if the sound becomes higher in pitch than if it becomes lower in pitch, and vice versa.

METHOD

If it is possible to attend selectively to the onset pitch of a stimulus while ignoring the direction of pitch change, then the time it takes to judge whether a tone is high or low in pitch, and the accuracy of this judgment, should not depend on pitch change. That is, whether a high pitch becomes lower or higher in pitch and a low pitch becomes lower or higher in pitch should not affect performance.

To investigate interactions between the dimensions of pitch and pitch change, twelve 250-ms pitch-glide stimuli were presented (see Figure 1 for the stimulus parameters). Sixty-four Rice University undergraduate subjects with normal hearing were instructed to listen to the sounds, attending selectively either to the pitch or to the direction of pitch change (depending on their experimental group), and then make speeded button press responses according to whether the pitch was high or low, or becoming higher or lower, respectively. The response buttons were arranged vertically with respect to one another.

Stimuli

Stimuli for which the sound started high and became higher or started low and became lower in pitch were considered "congruent", since their onset position (in the pitch space) and direction of pitch change corresponded, whereas the stimuli whose starting position and direction of pitch change were opposite were called "incongruent". The stimuli can also be described in terms of their difference from the average pitch of the set, with each stimulus being considered to have a small, medium or large "separation" (see Figure 1).

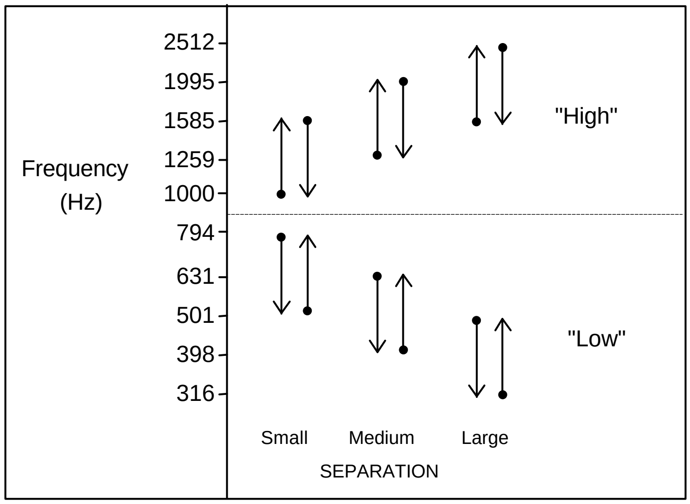

Figure 1. Schematic representation of the 12 auditory stimuli. The circle represents the starting pitch, and the arrow indicates the direction of pitch change, with the final pitch at the tip of the arrow. Note that the vertical axis is log frequency; thus, the equal arrow lengths reflect equal changes in perceived pitch for each stimulus. The horizontal line in the figure represents the relative "middle" of this pitch space: Stimuli that started above this frequency were considered relatively "high" in pitch and the others were considered relatively "low" in pitch. The stimuli are also labeled in terms of their separation, or relative distance from the middle of the stimulus set.

Subjects completed a block of 60 practice trials, then performed 2 blocks of 60 experimental trials, for a total of 180 trials in what was the first of two sessions (only Session 1 is reported here). There were five repetitions of each of the twelve stimuli randomly presented within each block. Accuracy feedback was given on each trial, and overall accuracy was presented at the end of each block. Half of the subjects responded to the onset pitch (or "position") of the stimulus (ignoring the direction of pitch change), and the other subjects responded to the direction of pitch change (ignoring onset position).

RESULTS

Mean correct response times (RTs) and mean accuracy were the dependent measures in separate ANOVAs (see [18], Experiment 2, for the complete ANOVA model and detailed results). RT results, as a function of cue dimension, separation and congruency, are presented in Figure 2. The main effects, two-way interactions and the three way Cue Dimension x Congruency x Separation were all significant. Accuracy results mirrored the RT results, so they will not be discussed in this report.

Overall, responses were faster for position judgments than for direction judgments [628 vs. 784 ms, respectively; $F(1, 56) = 12.14$, $p < .0010$, $MSE = 770{,}198$], indicating that listeners were faster to respond to the pitch of a stimulus than to changes in pitch.

Responses to congruent stimuli were faster than responses to incongruent stimuli [676 vs. 737 ms; $F(1, 56) = 47.48$, $p < .0001$, $MSE = 30{,}777$]. This congruency reflects that listeners failed to attend selectively to just one of the stimulus dimensions. Instead, the interaction of pitch and pitch change affected the relative speed and accuracy of responses.

As separation increased, overall RT decreased [mean RT = 728 ms at the smallest separation, 696 at the medium separation and 695 ms at the largest separation, $F(2, 112) = 16.69$, $p < .0001$, $MSE = 111{,}141$]. However, separation interacted with both congruency and cue dimension.

To understand how pitch and pitch change interact, it is necessary to understand how the three factors of cue dimension, separation and congruency interacted. This Cue Dimension x Congruency x Separation interaction (shown in Figure 2) was significant, $F(2, 112) = 30.54$, $p < .0001$.

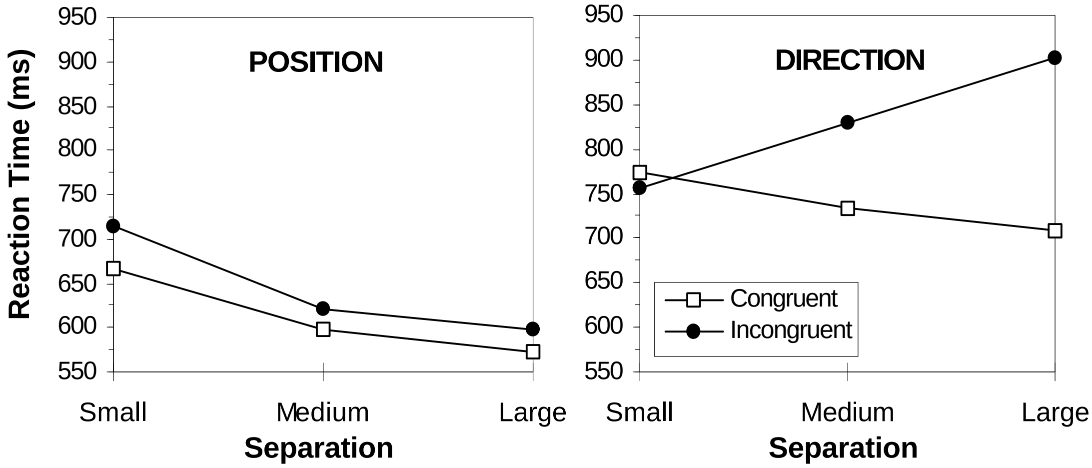

Figure 2. Mean RT results as a function of cue dimension, congruency and separation.

For subjects who attended to position (left side of Figure 2), responses to both congruent and incongruent stimuli were faster with increased stimulus separation (congruent: mean RT = 666, 597 and 573 ms; incongruent: mean RT = 715, 622, 598 ms for small, medium and large separations, respectively). It makes sense that more separated pitches would result in faster responses when pitch position is the cue dimension, since separation can also be considered a measure of stimulus discriminability for pitch. Within the stimulus set, it should be easiest to respond to stimuli that are more discriminable. The slowdown due to incongruent direction information is fairly constant with increased separation, suggesting that direction information has a uniform (and fairly small) effect on position judgments for all of the stimuli. Of course, it should be noted that the change in pitch was equal across stimuli.

When direction was the relevant dimension (right side of Figure 2), responses to congruent stimuli (open squares) were faster with increased separation (mean RT = 775, 734 and 709 ms), whereas responses to incongruent stimuli (filled circles) showed the opposite pattern, namely a pronounced slowing with increased stimulus separation (mean RT = 757, 830 and 902 ms for small, medium and large separations, respectively). Responses to congruent stimuli were, on the whole, still faster than responses to incongruent stimuli, but there was relatively little gain in performance with increased separation. That is, direction judgments seemed less aided by greater separation when the pitch and the direction matched. When the pitch was incongruent with the direction of pitch change, however, greater separation meant much slower responses to direction. Hence, it appears that the further the pitches were from the middle of the pitch space, the more salient pitch became, and the greater its interference with direction judgments. This was especially true for the largest level of separation. When separation was small, it was apparently harder to discriminate the different pitches, making pitch less salient. In this case it was apparently easier to ignore the irrelevant position information, as indicated by better performance on the direction judgment.

CONCLUSIONS

The dimensions of pitch and pitch change interacted, so that responses were almost always faster when the direction of pitch change matched the onset pitch position (i.e., for congruent stimuli). This parallels the results of studies using static auditory stimuli, as well as both static and dynamic visual stimuli. It is important to note that dynamic auditory stimuli (i.e., changing pitches) seem to behave in some ways similar to dynamic visual stimuli. That is, the dimension of position influences the perception of, and responding to, direction, and vice versa. This makes it clear that both pitch and pitch change need to be considered when designing auditory stimuli and auditory displays, but it also provides a visual analogy to help designers remember and understand the interaction of these two stimulus dimensions.

The intrusion of the information of one dimension onto judgments regarding the other dimension was not symmetrical, in that pitch information had a greater influence on responses to direction of pitch change than direction information had on pitch judgments. However, in both cases, when the two dimensions provided congruent information (e.g., a high pitch also became higher in pitch) responses were faster than if the two dimensions provided incongruent information (e.g., a high pitch became lower in pitch). Auditory display designers should take advantage of this congruency effect when a crucial distinction must be made between high and low pitches. Note that this proposed use of the congruency effect is conceptually similar to Kramer's [6] suggestion of mapping a single data stream to more than one sound parameter.

It was also observed that the congruency effect was fairly uniform for pitch judgments across the range of stimuli. However, for pitch-change judgments, responses to the stimuli that were more extreme in pitch (as compared to the average pitch of the set) were more influenced by irrelevant pitch information. Thus, these results suggest that for the design of auditory displays, the nature of the task must be considered in deciding upon the range of stimuli to use. In particular, if *pitch* is the dimension of a dynamic auditory message to which an operator must attend, there will always be some intrusion of the direction information. Selective attention is not perfect. However, a greater separation in the onset of pitches will mean better performance on the task, with no increased interference from direction information as a result. If pitch *change* is the important dimension, then in terms of performance on a selective listening task with an auditory display there is little to be gained with increased pitch separation in the case of congruent stimuli, and much to be lost in the case of incongruent stimuli. Hence, to reduce the deleterious effects of the irrelevant pitch information, an auditory display designer should restrict the range over which stimulus onsets may vary for a given task.

REFERENCES

1. Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.
2. Deutsch, D. (Ed.) (1982). *The psychology of music*. New York: Academic Press.
3. Ehrenstein, W. H. (1994). The Simon effect and visual motion. *Psychological Research, 56*, 163-169.
4. Grau, J. W., & Kemler Nelson, D. G. (1988). The distinction between integral and separable dimensions: Evidence for the integrality of pitch and loudness. *Journal of Experimental Psychology: General, 117*, 347-370.
5. Kramer, G. (1994). Some organizing principles for representing data with sound. In G. Kramer (Ed.) *Sonification, audification, and auditory interfaces*. Reading, MA: Addison Wesley.
6. Kramer, G. (1996). Mapping a single data stream to multiple auditory variables: A subjective approach to creating a compelling design. *Proceedings of The Third International Conference on Auditory Display,* (November 4-6) Palo Alto, CA.
7. Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance, 8(2)*, 177-193.
8. Melara, R. D., & Marks, L. E. (1990). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception and Performance, 16(2)*, 398-414.
9. Melara, R. D., Marks, L. E., & Lesko, K. E. (1992). Optional processes in similarity judgments. *Perception and Psychophysics, 51(2)*, 123-133.
10. Michaels, C. F. (1988). S-R compatibility between response position and destination of apparent motion. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 231-240.
11. Michaels, C. F. (1993). Destination compatibility, affordances, and coding rules: A reply to Proctor, Van Zandt, Lu, and Weeks. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 1121-1127.
12. Moore, B. C. J. (1989). *An introduction to the psychology of hearing (3rd Edition)*. London: Academic Press.
13. Proctor, R. W., Van Zandt, T., Lu, C.-H., & Weeks, D. J. (1993). Stimulus-response compatibility for moving stimuli: Perception of affordances or directional coding? *Journal of Experimental Psychology: Human Perception and Performance, 19*, 81-91.
14. Révész, G. (1954). *Introduction to the psychology of music*. Norman, Oklahoma: University of Oklahoma Press.
15. Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51*, 300-304.
16. Stevens, S. S. (1935). The relation of pitch to intensity. *Journal of the Acoustical Society of America*, 6, 150-154.
17. Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64, 153-181.
18. Walker, B. N. (1997). *Congruency effects with dynamic auditory stimuli*. Unpublished Masters Thesis. Rice University, Houston, Texas. Available at URL: http://www.owlnet.rice.edu/~walkerb/research/masters/
19. Walker, R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception and Psychophysics, 42(5)*, 491-502.