The effect of onset asynchrony on relative weights in profile analysis^{a)}

Jinyu Qian^{b)} and Virginia M. Richards^{c)}

Department of Psychology, University of Pennsylvania, 3401 Walnut Street, Suite 302C, Philadelphia, Pennsylvania 19104

(Received 9 September 2009; revised 17 January 2010; accepted 20 January 2010)

Decision weights were estimated in a profile analysis task to determine whether onset asynchronies between the signal component and the nonsignal components encourage the segregation of the signal relative to the other components. The signal component onset was either synchronous or asynchronous with respect to the nonsignal components. In the asynchronous conditions, thresholds were higher and the decision weights were less efficient than in the synchronous conditions. These data are largely consistent with a segregation hypothesis: onset asynchrony encourages subjects to shift strategies from one of spectral shape discrimination toward one of intensity discrimination.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3314251]

PACS number(s): 43.66.Fe, 43.66.Ba [MW]

Pages: 2461–2465

I. INTRODUCTION

Past research has shown that the ability to detect changes in spectral shape depends on the relative onsets of the components comprising the stimulus. Consider, for example, the detection of an increment to the level of the central component (signal component) of several equal-amplitude tones (nonsignal components). Thresholds increase when the onset of the signal components is changed from synchronous to asynchronous relative to the nonsignal components (Green and Dai, 1992; Hill and Bailey, 1997; Lentz *et al.*, 2004). One potential account for this result is the segregation hypothesis: the onset asynchrony between the signal and nonsignal components results in the segregation of the signal from the nonsignal components. Consequently, there is no longer a common reference system for the two segregated sounds, and thus level comparisons between the signal and nonsignal components are compromised. An implication of this hypothesis is that as the signal and nonsignal components become segregated, the subjects' decision strategy would shift from one associated with changes in spectral shape to one associated with changes in intensity at the signal frequency.

To provide an example, Hill and Bailey (1997) found that thresholds were higher when the signal component preceded the nonsignal components by 320 ms compared to when all components shared a common onset. To reduce the degree of segregation, they introduced two "captor tones," one higher and one lower in frequency than the signal frequency. The captor tones were turned on with the signal and were turned off when the nonsignal components were turned on. Thus the early portion of the signal could be grouped

with the captor tones, freeing the latter portion of the signal component to be grouped with the nonsignal components. Consistent with this argument, the captor tones did result in lower thresholds relative to when there were no captor tones. Not all methods to promote the segregation of the signal component and thresholds in profile analysis, however, have succeeded. Hill and Bailey (2000) examined the effect of mistuning the signal component relative to the harmonically related nonsignal components. Those measures failed to indicate a change in threshold when the signal component was perceptually segregated from the nonsignal components.

The aim of the current study is to examine the effect of onset asynchrony on decision strategies in a profile analysis task. The subjects' task is to detect which of two sounds has an increment to the 1000 Hz component relative to the other equal-amplitude components. Decision strategies are revealed by estimating decision weights, or relative weights, for conditions in which the signal onset is either asynchronous or synchronous relative to the nonsignal components. Both 5- or 15-tone complexes are tested in different sets of conditions to investigate whether the effect of onset asynchrony on relative weights varies depending on the number of frequency components. In addition, the amount of asynchrony is varied in different conditions to reveal systematic changes in the results.

In each condition, relative weights for each frequency component and for the overall level are estimated (Berg and Green, 1990; Berg, 2004). For an ideal (linear) profile analysis observer, if the relative weight for the signal component is arbitrarily set to 1, the weight for the nonsignal components should be $-1/(n-1)$, where n is the total number of frequency components of the multitone complex (Berg and Green, 1990). Additionally, for an ideal profile-listening strategy, there should be no effect of level randomization (i.e., a weight of zero associated with overall level randomization). In contrast, a strict intensity discrimination decision strategy would not incorporate information from the nonsignal components and would show a strong effect of overall

^{a)} Portions of this research were presented at the 31st MidWinter Meeting of Association for Research in Otolaryngology, Phoenix, AZ, 16–21 February 2008, Abstract 911.

^{b)} Author to whom correspondence should be addressed. Electronic mail: qjyspace@yahoo.com

^{c)} Present address: Department of Cognitive Sciences, University of California, 3151 Social Science Plaza, Irvine, CA, 92697.

level randomization. Overall, if subjects' detection strategies change from an across-frequency comparison strategy to an intensity discrimination strategy when an onset asynchrony is introduced, two related predictions can be made: (a) as the onset asynchrony increases, the relative weights associated with overall level randomization should increase, and (b) the weights associated with the nonsignal frequencies should approach zero. If, however, onset asynchrony does not influence decision strategies but affects processes such as the efficiency with which the subjects' decision rules are carried out (e.g., an overall increase in "internal noise"), no changes in the pattern of relative weights would be expected. To discriminate between these alternatives, in the current experiment subjects' efficiencies as well as decision weights are estimated (see [Berg, 2004](#)).

II. METHODS

A. Stimuli

The standard stimulus consisted of equal-amplitude tones equally spaced on a logarithmic frequency scale with frequencies ranging from 200 to 5000 Hz. In the target stimulus, a 1000 Hz tone was added in phase to the 1000 Hz signal component of the standard stimuli. The phase of each component was randomly chosen from a uniform distribution for each trial but fixed across two intervals. In order to derive relative weights for all components, independent level perturbations drawn from a standard normal distribution ($\mu = 0$ dB, $\sigma = 1$ dB) were added in phase to each component in each interval. On each presentation, the overall level of the stimulus was drawn at random from a uniform distribution ranging from 40 to 60 dB sound pressure level.

There were six conditions: the 5- or 15-tone simultaneous conditions, the 5- or 15-tone leading-100 conditions, and the 5- or 15-tone leading-300 conditions. In the first two conditions, the signal and nonsignal components had the same onset. In the last four leading conditions, the signal component started earlier than the nonsignal components by either 100 or 300 ms. The multitone stimuli in each condition had either 5 or 15 frequency components. All nonsignal components had the same onset and the same duration of 200 ms. All stimuli had 5 ms raised cosine rise/fall ramps at the stimulus onsets and offsets.

The stimuli were presented diotically over Sennheiser HD410 SL headphones with visual feedback following each of the subjects' responses. Subjects were tested in a double-walled sound-attenuated booth in a quiet room and the stimuli were generated by TDT system II hardware controlled by MATLAB software.

B. Subjects and procedure

Four young adults aged 20–31 with normal hearing (<15 dB hearing loss at frequencies of 250, 500, 1000, 2000, 4000, 6000, and 8000 Hz in both ears) served as subjects. A two-alternative forced-choice procedure with a 2-down 1-up adaptive rule was used to estimate the threshold corresponding to 71% correct responses ([Levitt, 1971](#)) or a d' of 0.77. Thresholds and relative weight estimates are based on 50 50-trial runs (a total of 2500 trials) for each

condition. Subjects S1 and S2 ran the 5-tone simultaneous and the 5-tone leading-300 conditions in different orders, followed by the 15-tone simultaneous and the 15-tone leading-300 conditions in different orders, and then the 5- and 15-tone leading-100 conditions in different orders. Subjects S3 and S4 ran the three 15-tone conditions in different orders, followed by the three 5-tone conditions in different orders. All subjects had at least four hours of practice prior to data collection.

C. Relative weight estimation

1. Decision model

For each trial the decision variable (DV) was modeled as the difference between the weighted sums of the level of each frequency component in the first interval and in the second interval ([Berg, 2004](#)).¹ The specific weights assigned to different frequency components are the relative weights to be estimated. Theoretically, the sum of the relative weights for individual frequency components forms the weight for overall level randomization ([Berg, 2004](#)). However, to avoid the cumulative error of estimation, the weight for overall level randomization was estimated independently.

2. Relative weights

Logistic regression was applied to estimate relative weights (e.g., [Alexander and Lutfi, 2004](#); [Dye et al., 2005](#)) for the overall level randomization and each frequency component of the multitone stimuli. Using a MATLAB build-in function (`glmfit`), the relative weights and the standard errors of these weights were estimated from a predictor matrix and a response array. The predictor matrix included the differences in the overall level randomization between two intervals for every trial and the level differences for each frequency component between the two intervals excluding the differences in the overall level randomization. The response array consisted of the responses coded as zeros ("target in interval 2") or ones ("target in interval 1"). The relative weights were based on 2500 trials (50 sets of 50 trials) for each condition. Each set of relative weights was normalized so that the weight of the signal component was 1. The standard errors of the relative weights were normalized accordingly.

3. Measures of efficiency

Efficiency measures allow a comparison between the ideal and subjects' performance and between different subjects' performance. The ideal observer maximizes the d' based on the best linear decision model. Here, overall efficiency, η , is described as the product of two quantities: the weighting efficiency, η_w , and the noise efficiency, η_n . ([Berg, 1990](#); [Doherty and Lutfi, 1999](#); [Berg, 2004](#)). The weighting efficiency describes the efficiency of the subject's weighting strategy relative to the ideal weighting strategy. The noise efficiency describes the efficiency associated with processes other than the pattern of relative weights alone, such as peripheral and central internal noises. Following the notation of [Berg \(2004\)](#), the three efficiencies are defined as follows:

TABLE I. Individual and averaged thresholds (dB: signal level relative to the standard level) are shown. The standard errors of the mean are shown at the bottom row.

Subjects	Conditions					
	5-tone			15-tone		
	Sim	Lead-100	Lead-300	Sim	Lead-100	Lead-300
S1	-9.5	-6.7	-2.0	-14.8	-5.3	-1.3
S2	-6.1	1.4	0.6	-12.8	-4.0	0.1
S3	-11.0	-4.1	-1.8	-6.7	-2.0	0.3
S4	-12.0	-0.4	1.6	-13.1	-5.5	-2.7
AVG	-9.7	-2.4	-0.4	-11.8	-4.2	-0.9
SEM	1.3	1.8	0.9	1.8	0.8	0.7

$$\eta = \eta_W \times \eta_N = (d'_W/d')^2 \times (d'_{\text{obs}}/d'_W)^2, \quad (1)$$

where d' is the sensitivity index for an ideal observer, d'_W is the highest sensitivity index based on the estimated pattern of relative weights, and d'_{obs} is the subject's actual sensitivity index.

The d'_W calculation was based on the Eq. (2) below given by Berg and Green (1990) and Berg (2004).³ For each subject this equation was applied using the signal level appropriate for the various conditions,

$$d'_W = \frac{\sqrt{2\Delta_s}}{\sqrt{\sum_{i=1}^n w_i^2 \sigma^2 + w_L^2 \sigma_L^2}}, \quad (2)$$

where Δ_s is the signal level increment in dB at the subjects' thresholds, w_i and w_L are the relative weights for the i th frequency component and the weight for overall level randomization, respectively. The variance of the level perturbations is given by σ^2 and σ_L^2 is the variance of overall level randomization. The number of frequency components is given by n .

The same equation was used to calculate the d' , the sensitivity index for an ideal observer, except that the subjects' relative weights were replaced with the ideal relative weights in Eq. (2).

III. RESULTS AND DISCUSSION

A. Psychophysical results

Table I shows the thresholds (in dB signal re. standard) for the four subjects and the averaged data. In the simultaneous conditions, thresholds are lower for the 15-ton than the 5-tone condition, although the difference is only 2 dB. A two-way, within-subjects analysis of variance (ANOVA) showed no significant difference between the 5- and 15-tone conditions, and no significant interaction between the number of components and onset asynchrony. In contrast, the effect of onset asynchrony was significant [$F(2,6)=92$, $p < 0.0005$]. These data are in accord with past results, except that in the current experiment the change in thresholds with number of components is somewhat smaller than in the past (Green and Dai, 1992; Hill and Bailey, 1997; Lentz *et al.*, 2004; Berg, 2004).

FIG. 1. (Color online) The values of relative weight for level randomization averaged across subjects. Each bar represents one condition and error bars indicate 1 standard error of the mean. For each onset asynchrony, the left and the right bars are for the 5-tone and 15-tone stimuli, respectively.

B. Subject's relative weights

Figure 1 shows the values of relative weights for level randomization averaged across subjects for the six conditions tested (from left to right: 5-tone and 15-tone simultaneous, 5-tone and 15-tone leading-100, and 5-tone and 15-tone leading-300). Error bars show the standard errors of the mean. The results for individual subjects are similar to the pattern of results shown in Fig. 1, with the exception that for S4 in the 15-tone condition the relative weights are approximately the same across the three levels of onset asynchrony. A two-way, within-subjects ANOVA indicated a significant effect of onset asynchrony [$F(2,6)=28.8$; $p < 0.001$] but did not indicate a significant effect of number of components⁴ nor a significant interaction between number of components and onset asynchrony. These results are consistent with the segregation argument, suggesting that the signal component may be perceptually segregated from the other components in the leading conditions and that the subjects may, at least partially, shift their strategies from the spectral shape discrimination in the simultaneous conditions to intensity discrimination in the leading conditions.

Normalized relative weights for all of the components are plotted as a function of frequency in Fig. 2(a) (5-tone conditions) and Fig. 2(b) (15-tone conditions). Each panel includes one subject's data from simultaneous (circles), leading-100 (diamonds) and leading-300 (squares) conditions. The standard errors of the estimated relative weights ranged from approximately 0.01 to 0.25 across subjects and conditions (not shown in Fig. 2).

According to the second prediction of the segregation hypothesis, one would expect the relative weights at the non-signal frequencies to approach zero as the signal onset asynchrony increases. For 5-tone conditions shown in Fig. 2(a), nonsignal weights at frequencies lower than the signal frequency are negative for simultaneous (circles) conditions but approach zero for the leading-300 conditions (squares). However, the relative weights for components at frequencies higher than the signal frequency are approximately zero in all conditions.

For the 15-tone stimuli shown in Fig. 2(b), the nonsignal

FIG. 2. (Color online) Relative weights plotted as a function of frequency are shown for each subject (panels) for the (a) 5-tone complex and (b) 15-tone complex stimuli. Values are plotted separately for the simultaneous (circles), leading-100 (diamonds), and leading-300 (squares) conditions.

weights are near zero and do not appear to systematically change as the asynchrony grows. This is not surprising—for the simultaneous conditions normalized ideal nonsignal weights are $-1/4$ for the 5-tone stimuli but $-1/14$ for the 15-tone stimuli.

In an effort to provide an overview of the change in relative weights with onset asynchrony, and because we had no *a priori* hypotheses regarding changes in the complex pattern of relative weights, for each of the three onset asynchronies (0, 100, and 300 ms), the relative weights were averaged across all nonsignal components. Then a within-subjects two-factor ANOVA was run on these averages to determine whether the nonsignal relative weights gravitate toward zero as the onset asynchrony increases. The results did not reveal a significant effect of number of components,⁴ nor was the interaction between number of components and onset asynchrony significant. The main effect of onset asynchrony, however, was significant [$F(2,6)=55.3$, $p < 0.001$]. For the 5-tone stimuli, the relative weights averaged across subjects and all nonsignal components were -0.17 , -0.15 , and 0.00 for the simultaneous, leading-100, and leading-300 conditions, respectively. As one would anticipate, for the 15-component conditions, the change was smaller: -0.05 , -0.05 , and 0.02 for the simultaneous, leading-100, and leading-300 conditions, respectively. Overall, these results are in accord with the argument that onset asynchrony makes across-frequency comparisons of level more difficult, leading to a shift in strategy.

To summarize, the relative weights are consistent with a segregation-based argument: as the signal onset asynchrony is increased, subjects shift from a spectral shape discrimination strategy toward an intensity discrimination strategy because across-frequency level comparison of the signal and nonsignal components becomes difficult. First, as the onset asynchrony increases, the relative weight associated with the overall level randomization increases. Second, as the signal onset asynchrony increases, the average of the relative weights of the nonsignal components approaches zero. Both of these factors suggest a shift from detecting differences in spectral shape toward detecting changes in intensity at the signal frequency.

C. Weighting efficiency and noise efficiency

The weighting and noise efficiencies, for the individual subjects and their averages, are listed in Tables II and III, respectively. Consistent with the pattern of relative weights described above, the weighting efficiencies fall dramatically as the onset asynchrony increases for both the 5-tone and 15-tone stimuli. The noise efficiencies grow somewhat as the onset asynchrony increases for the 5-tone stimuli but do not appear to change for the 15-tone stimuli. The results of within-subjects two-way ANOVAs are consistent with this observation. For the weighting efficiencies, the effect of onset asynchrony is significant [$F(2,6)=38$, $p < 0.0001$] but the number of components, and the interaction term, did not approach significance. For the noise efficiencies, neither main effect nor the interaction reached significance. The factor of onset asynchrony, did, however, approach significance [$F(2,6)=3.3$, $p \sim 0.1$]. Overall, the effect of asynchrony on weighting efficiencies far exceeded the effect on noise efficiencies, suggesting that threshold elevations due to onset asynchrony reflect “inefficient” changes in decision strategies rather than increases in internal noise.

IV. SUMMARY AND CONCLUSIONS

Spectral shape discrimination deteriorates when an onset asynchrony is introduced between signal and nonsignal fre-

TABLE II. Individual and averaged weighting efficiencies are shown. The standard errors of the mean are shown at the bottom row.

Subjects	Conditions					
	5-tone			15-tone		
	Sim	Lead-100	Lead-300	Sim	Lead-100	Lead-300
S1	0.24	0.09	0.03	0.55	0.10	0.03
S2	0.11	0.03	0.02	0.43	0.25	0.03
S3	0.42	0.07	0.04	0.25	0.04	0.03
S4	0.58	0.17	0.03	0.48	0.14	0.11
AVG	0.34	0.09	0.03	0.43	0.13	0.05
SEM	0.10	0.03	0.00	0.06	0.04	0.02

TABLE III. Individual and averaged noise efficiencies are shown. The standard errors of the mean are shown at the bottom row.

Subjects	Conditions					
	5-tone			15-tone		
	Sim	Lead-100	Lead-300	Sim	Lead-100	Lead-300
S1	0.25	0.36	0.41	0.28	0.22	0.33
S2	0.28	0.29	0.38	0.23	0.07	0.31
S3	0.19	0.29	0.35	0.12	0.28	0.29
S4	0.17	0.07	0.24	0.22	0.17	0.13
AVG	0.22	0.25	0.35	0.21	0.19	0.27
SEM	0.03	0.06	0.04	0.03	0.04	0.05

quency components of the stimuli (Green and Dai, 1992; Hill and Bailey, 1997; Lentz *et al.*, 2004). It has been suggested that asynchrony leads to the perceptual segregation of the signal component relative to the nonsignal components, impairing the across-frequency comparison required for spectral shape discrimination.

Relative weights were derived to explore the effect of onset asynchrony on decision strategies in profile analysis. Across subjects and conditions, both thresholds and the relative weights for the overall level randomization increase with increases in onset asynchrony. Moreover, the average of the relative weights for the nonsignal components approach zero as the onset asynchrony increases. These features are consistent with expectations that increasing the onset asynchrony hinders across-frequency comparisons of level, forcing subjects to shift from a strategy of discriminating spectral shape toward a strategy of discriminating changes in level at the signal frequency (intensity discrimination). The measurement of subjects' noise efficiencies do not provide support for an alternative account—that thresholds increase with increases in onset asynchrony because the internal noise is effectively increased. Overall, the current results support the segregation hypothesis: as the onset asynchrony between the signal and the other components increases, thresholds deteriorate because subjects can no longer compare the levels of the signal component and the remaining components.

ACKNOWLEDGMENTS

We would like to thank the associate editor, Dr. Hedwig E. Gockel, and Dr. Robert A. Lutfi who provided valuable comments on early versions of the paper. This work was

supported by Grant No. R01 DC002012 from the National Institutes of Health.

¹The decision variable, DV, is defined as $DV = \sum_{i=1}^n w_i \Delta I_i$, where the w_i are the relative weights for i th frequency component, and the ΔI_i are the level differences between the two intervals of the i th frequency component. Note that ΔI_i is the sum of the overall level difference, ΔL , and the level difference excluding the overall level difference, Δp_i . Thus, the decision variable may be rewritten as $DV = \sum_{i=1}^n w_i (\Delta L + \Delta p_i) = \Delta L \sum_{i=1}^n w_i + \sum_{i=1}^n w_i \Delta p_i = \Delta L w_L + \sum_{i=1}^n w_i \Delta p_i$, where $w_L = \sum_{i=1}^n w_i$.

³The development of Eq. (2) assumed normally distributed level variation in addition to normally distributed perturbations. Here uniformly distributed level variation was tested. Simulations were also run to provide estimates of d'_w and subsequent values of efficiency (e.g., Qian and Richards, 2008). The results were not appreciably different from those obtained using Eq. (2).

⁴A main effect of the number of components, 5 vs. 15, provides little information because the relative weights (a) have been independently normalized and (b) are not assured to have the same scaled values prior to normalization. This factor is included in the ANOVA to allow for a test of an interaction.

- Alexander, J. M., and Lutfi, R. A. (2004). "Informational masking in hearing-impaired and normal-hearing listeners: Sensation level and decision weights," *J. Acoust. Soc. Am.* **116**, 2234–2247.
- Berg, B. G. (1990). "Observer efficiency and weights in a multiple observation task," *J. Acoust. Soc. Am.* **88**, 149–158.
- Berg, B. G. (2004). "A molecular description of profile analysis: Decision weights and internal noise," *J. Acoust. Soc. Am.* **115**, 822–829.
- Berg, B. G., and Green, D. M. (1990). "Spectral weights in profile listening," *J. Acoust. Soc. Am.* **88**, 758–766.
- Doherty, K. A., and Lutfi, R. A. (1999). "Level discrimination of single tones in a multitone complex by normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **105**, 1831–1840.
- Dye, R. H., Stellmack, M. A., and Jurcin, N. F. (2005). "Observer weighting strategies in interaural time-difference discrimination and monaural level discrimination for a multi-tone complex," *J. Acoust. Soc. Am.* **117**, 3079–3090.
- Green, D. M., and Dai, H., (1992). "Temporal relations in profile comparisons," in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner, (Pergamon, Oxford), pp. 471–477.
- Hill, N. I., and Bailey, P. J. (1997). "Profile analysis with an asynchronous target: Evidence for auditory grouping," *J. Acoust. Soc. Am.* **102**, 477–481.
- Hill, N. I., and Bailey, P. J. (2000). "Profile analysis of harmonic complexes: Effects of mistuning the target," *J. Acoust. Soc. Am.* **107**, 2291–2294.
- Lentz, J. J., Leek, M. R., and Molis, M. R. (2004). "The effect of onset asynchrony on profile analysis by normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **116**, 2289–2297.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Qian, J., and Richards, V. M. (2008). "Relative weights for frequency glide detection using narrowband noise," *J. Acoust. Soc. Am.* **123**, 2692–2700.