

Spectral processing of two concurrent harmonic complexes

Yi Shen^{a)} and Virginia M. Richards

Department of Cognitive Sciences, University of California, Irvine, California 92697-5100

(Received 7 April 2011; revised 26 October 2011; accepted 26 October 2011)

In a concurrent profile analysis task, each of the two observation intervals was the sum of two harmonic complexes. In the first interval one of the harmonic complexes had a flat spectrum and the other had a broad spectral peak at 1 kHz. In the second interval, the association between the spectral profiles and the complexes was either consistent with the first interval, or inconsistent so that profile changes (flat versus peaked) could be created in both of the complexes. In two experiments, thresholds and psychometric functions for detecting the profile change were measured in terms of the spectral peak's magnitude as functions of three types of segregation cues: Difference in fundamental frequency, onset asynchrony, and difference in interaural time difference between the two complexes. Decreasing the magnitude of each cue led to higher thresholds, and shallower psychometric functions whose upper asymptotes often failed to reach 100% correct. The patterns of the threshold and psychometric functions varied across cue types and across individual listeners. The results suggest that informational masking is present in the concurrent profile analysis task. Segregation cues appear to contribute to the release from informational masking, but the process depends on listening strategies adopted by individual listeners.

© 2012 Acoustical Society of America. [DOI: 10.1121/1.3664081]

PACS number(s): 43.66.Jh, 43.66.Dc [EB]

Pages: 386–397

I. INTRODUCTION

Humans have the extraordinary ability of comprehending complex acoustic environments. For normal-hearing listeners, the ability to segregate a target sound source in an interfering background, or tell musical instruments apart in a symphonic setting, appears effortless. This phenomenon, sometimes referred to as auditory scene analysis (e.g., Bregman, 1990), has been studied extensively through psychophysical and physiological approaches. Moreover, the acoustic cues that promote concurrent source segregation have been identified (for a recent review, see Micheyl and Oxenham, 2010).

The goal of the present study is to develop a behavioral experimental paradigm to study how segregation cues affect the perception of the spectral envelopes (or spectral profiles) of two concurrent sound sources. Previous studies addressing the role of segregation cues in concurrent-spectral processing tasks have used the double-vowel paradigm. For this task, listeners identify simultaneously presented vowel pairs (e.g., Zwicker, 1984; Assmann and Summerfield, 1989, 1990; Meddis and Hewitt, 1992; Culling and Darwin, 1993; de Cheveigné *et al.*, 1997; Lentz and Marsh, 2006; Hedrick and Madix, 2009). The listeners' ability to correctly identify both concurrent vowels typically improve when segregation cues are introduced. Assmann and Summerfield (1990) measured subjects' percent-correct identification scores as a function of Δf_0 , the difference in fundamental frequency across the simultaneous vowel pairs. For a stimulus duration of 200 ms, they found that introducing a Δf_0 could improve the percent correct by as much as 20 percentage points. Performance improved as Δf_0 increased from 0 to 1 semitone, and tended

to plateau for Δf_0 's greater than 1 semitone. Another strong cue for auditory source segregation is onset asynchrony (e.g., Darwin, 1984; Hukin and Darwin, 1995; Darwin and Hukin, 1997). Lentz and Marsh (2006) measured double-vowel identification in normal-hearing and hearing-impaired listeners with values of onset asynchrony ranging between 100 and 300 ms. Both groups of listeners showed better performance for larger onset asynchrony, a result that held even when the two vowels had the same f_0 . A third segregation cue is the difference in interaural time difference among concurrent sources (Δ ITD). Interaural time difference (ITD) is the dominant acoustic cue for azimuthal localization at frequencies below 1500 Hz (e.g., Sandel *et al.*, 1955). Presenting sound sources at different ITDs benefits the segregation of the sources, but the amount of the benefit is relatively small compared to the Δf_0 -based improvements (e.g., Shackleton and Meddis, 1992).

When two sound sources with overlapping spectra are presented simultaneously, like the stimuli used in the double-vowel experiments, the recognition of their individual spectral profiles could be undermined by several factors including (1) the two sounds might lead to overlapping excitation in the auditory periphery and energetically mask each other; (2) if the task depends on just one of the two sounds, the other sound could act as an interferer and increases the apparent internal noise due to incomplete segregation; and (3) they might be perceived as a single auditory object or two indistinguishable objects due to a failure of segregation. If we define any degradation in performance that cannot be accounted for by energetic masking as informational masking (e.g., Durlach *et al.*, 2003), both factors (2) and (3) mentioned previously could give rise to informational masking. Although the double-vowel paradigm provides an important tool to study the benefits of segregation cues on spectral processing, the existing data do not indicate whether the

^{a)}Author to whom correspondence should be addressed. Electronic mail: shen.yi@uci.edu

improvement in performance associated with the introduction of segregation cues is due to a release from energetic or informational masking. Because informational masking could potentially play an important role in the perception of concurrent sources (e.g., Brungart, 2001), alternative experimental techniques are necessary in order to address the roles of energetic and informational masking in competing-source perception, allowing a better description of the effect of segregation cues on sound source detection and identification.

One experimental technique to investigate features of masking is to measure psychometric functions (see, e.g., Lutfi *et al.*, 2003). Because double-vowel experiments are designed as identification tasks, the results typically describe only one point on the psychometric function. To provide a full view of the function, a detection or discrimination task using nonspeech stimuli may be preferable. A group of experiments that have demonstrated the promise of such an approach is the measurement of pitch discrimination for a target harmonic complex presented concurrently with a masker complex. Both Δf_0 and onset asynchrony have been shown to benefit the pitch discrimination, similar to the findings of the double-vowel experiments (see, e.g., Rasch, 1978; Carlyon, 1996a,b, 1997; Micheyl *et al.*, 2006; Bernstein and Oxenham, 2008). Although pitch-discrimination paradigm provides a useful way of assessing the sound-segregation capability of the auditory system, this method does not generalize to tasks measuring the perception of spectral envelopes (or formant structures).

In the current study, we introduce a new experimental method for investigating the influences of auditory source segregation on the perception of spectral envelope using a profile-analysis paradigm. We refer to the task as concurrent profile analysis. In traditional profile-analysis experiments (see, e.g., Green, 1983; Green and Kidd, 1983), listeners detect changes in the power spectrum of a stimulus, usually a complex composed of several tones. Although a few studies have investigated the effect of segregation cues on profile analysis (see, e.g., Hill and Bailey, 1997, 2000; Qian and Richards, 2010), they do not directly address listeners' sensitivity to the spectra of two competing sounds.

In a concurrent profile analysis task, listeners detect changes in power spectra of two concurrent harmonic complexes across two test intervals. One flat and one peaked spectral profile are presented in each trial, they are assigned to either the same (in a *same* trial) or different (in a *different* trial) complexes across the two observation intervals (see Fig. 1). If the two complexes are perceived as a single source, the profile of the mixture is not enough to differentiate *same* and *different* trials.¹ If the two complexes are perceptually segregated, the profile changes across intervals for both complexes in the *different* trials but not in the *same* trials. Therefore, segregation of the two complexes is required for above-chance performance in the concurrent profile analysis task. Within each trial, the two complexes might differ in terms of ITD, onset time, or fundamental frequency. These three types of acoustic cues are expected to help segregate the two complexes and improve performance.

As in traditional profile-analysis experiments, thresholds and psychometric functions can be obtained for concurrent

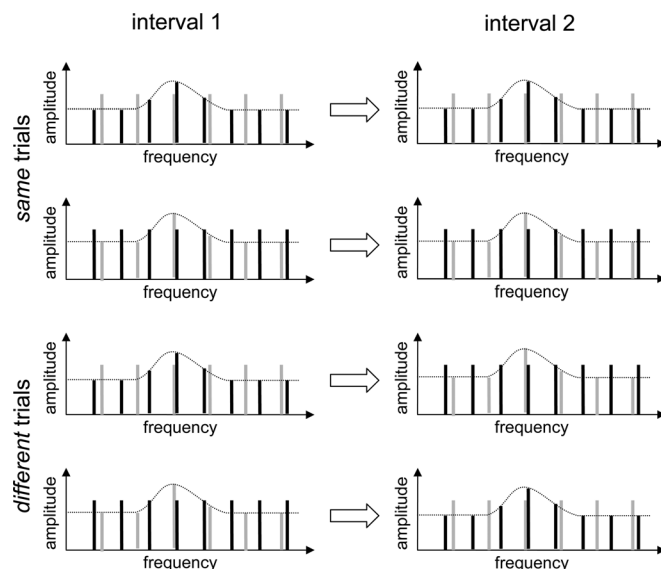


FIG. 1. Schematics of stimulus spectra in *same* and *different* trials. Different types of trials are arranged in rows, and the two stimulus intervals are arranged in columns. Within each panel, the dark and gray lines plot the spectra of the two concurrent complexes. The imposed spectral profile (a peak in the spectral envelope) is indicated by the dotted curve. The level randomization, implemented in the actual experiment, is not included here to enable clearer visualization.

profile analysis. For the concurrent profile analysis method, thresholds reflect the sensitivity to spectral profiles of concurrent sound sources, and psychometric functions describe how performance improves as the spectral profiles become more and more distinct. The current study measures both thresholds and psychometric functions for concurrent profile analysis in separate experiments. By studying thresholds and psychometric functions jointly as functions of Δf_0 , onset asynchrony, and Δ ITD between the two concurrent harmonic complexes, the relative importance of these acoustic cues in concurrent spectral-envelope processing can be better understood.

II. EXPERIMENT I: CONCURRENT PROFILE ANALYSIS—THRESHOLD DATA

Experiment I studies the effect of segregation cues (Δf_0 , onset asynchrony, and Δ ITD) on concurrent profile analysis thresholds. These thresholds reflect listeners' spectral-processing capabilities under concurrent stimulus presentation.

A. Methods

1. Stimuli

In this experiment, each of the two observation intervals in a trial consisted of two concurrent harmonic complexes. The two complexes differed in their fundamental frequencies, f_0 's, therefore one would sound lower in pitch than the other. Accordingly, the complexes will be referred to as the low- and high-pitch complexes. The complexes were generated by summing equal-amplitude tones with frequencies that are integer multiples of their fundamental frequencies. All harmonics below 4 kHz were included. The starting phase of each spectral component was drawn at random from a uniform distribution between 0 and 2π , separately for each complex. A

spectral peak at 1 kHz was introduced to either the low- or high-pitch complex, so that one complex had a “flat” spectral profile and the other had a “peaked” spectral profile.

The peaked spectral profile was generated by incrementing the amplitudes of harmonics within a two-octave range geometrically centered at 1 kHz. The amplitude increment (ΔA_i) on the i th harmonic depended on the absolute frequency of the component (f_i), and was given by

$$\Delta A_i = \begin{cases} a \left[1 - \cos \left(\frac{2(\ln f_i - \ln f_L)\pi}{\ln f_H - \ln f_L} \right) \right], & f_L \leq f_i < f_H \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where f_L and f_H (500 and 1000 Hz, respectively) are the lower and upper limits of the spectral peak and a is a positive scale factor whose value is varied to change the overall magnitude of the spectral increment. The number of components that have nonzero increments, N , depends on the complex’s fundamental frequency f_0 . If the peak amplitude of each component of a complex before introducing the increment is A , the amount of spectral change is quantified as the profile strength (PS) given by (Green *et al.*, 1987)

$$\text{PS} = 10 \log \frac{\sum \Delta A_i^2}{NA^2}. \quad (2)$$

During the stimulus generation, a desired PS value was obtained by setting the scale factor a in Eq. (1) to the appropriate value.

For each trial, the two stimulus intervals were 0.5 s in duration and were separated by a 0.5 s interstimulus silent pause. Whether the low- or high-pitch complex carried the peaked spectral profile was determined randomly for each stimulus interval. Thus, there were four distinct trial types, as illustrated in the four rows of Fig. 1: (1) In both intervals, the peaked profile was imposed onto the low-pitch complex. (2) The peaked profile was assigned to the high-pitch complex in both intervals. (3) In the first interval, the spectral increment was added to the low-pitch complex, whereas in the second interval, the spectral increment was applied to the high-pitch complex. (4) The spectral increment shifted from the high-pitch complex in the first interval to the low-pitch complex in the second interval. We can group cases (1) and (2) together and label them as the *same* trials, because of their consistent associations between f_0 ’s and spectral profiles. Similarly, cases (3) and (4) will be referred to as the *different* trials.

A concurrent profile analysis threshold was defined as the magnitude of PS required for the listener to correctly identify the *different* trials from the *same* trials (at 70.7% correct, see Sec. II A 3 for further details). For each listener, concurrent profile analysis thresholds were measured for various Δf_0 ’s, onset asynchrony, and Δ ITDs between the two complexes.

The f_0 differences tested were 0.5, 0.75, 1, 1.5, and 2 semitones. For each trial, the f_0 ’s of the low- and high-pitch complex were consistent across the two observation inter-

vals, and they were chosen in the following manner. First, one f_0 was drawn from a uniform distribution between 200 and 400 Hz. Then, it was randomly assigned to be the f_0 of either the low- or high-pitch complex. When it was assigned to the low-pitch complex, the f_0 of the high-pitch complex was obtained by incrementing the first f_0 by an appropriate number of semitones. On the other hand, if the first f_0 was assigned to the high-pitch complex, the f_0 of the low-pitch complex was calculated as a decrement to the first f_0 . This procedure was implemented to prevent listeners from relying on only one or two highly salient spectral components throughout the experiment due to fixed f_0 ’s. It is worth pointing out that although the f_0 was randomized across trials, the frequency of the profile peak was fixed at 1 kHz.

Onset asynchrony of 0, 20, 40, 80, and 160 ms were tested. For each trial, whether the low- or high-pitch complex was gated on earlier in time was consistent across the two intervals, and the leading and lagging roles were assigned to the low- and high-pitch complexes at random. Both complexes were gated off simultaneously, and 50 ms raised cosine onset and offset ramps were applied to each complex. The lagging complex had a fixed duration of 0.5 s. The duration of the leading complex, as well as the duration of the combined complex pair, was then 0.5 s plus the onset delay.

Two ITD differences were tested. In the 0-ITD conditions, both complexes were presented diotically, yielding a Δ ITD of 0 μ s. In the ± 400 -ITD conditions, ITDs of 400 μ s were applied to both complexes in opposite directions (Δ ITD = 800 μ s). That is, for one complex the left channel led the right channel by 400 μ s (ITD = 400 μ s) and it was lateralized near the left ear, whereas for the other complex the right channel led the left channel (ITD = -400 μ s) yielding a sound lateralized near the right ear.² For each trial, the positive and negative ITDs were assigned to the high- and low-pitch complexes at random. The 800- μ s Δ ITD was introduced using the following procedure: First, the high- and low-pitch complexes were generated. To apply an ITD of 400 μ s (left-leading) onto one of the complexes, the original stimulus was shifted 200 μ s earlier in time and sent to the left output channel; at the same time, it was delayed by 200 μ s and sent to the right output channel.³ An ITD of -400 μ s (right-leading) was generated using time shifts in opposite directions. For each channel, the two complexes with appropriate ITDs were then summed together and presented to the listener.

Various randomization processes were implemented in the current experiment. For a given experimental condition, i.e., a particular combination of Δf_0 , onset asynchrony, and Δ ITD, whether one complex in the concurrent complex pair had higher or lower f_0 , leading or lagging onset, left or right lateralization in each trial was determined randomly and independently. Importantly, within each trial, the two stimulus intervals shared the same pair of f_0 ’s, onset delays, and ITDs.

The overall levels of each complex in each of the two intervals were independently drawn from a uniform distribution between 50 and 70 dB sound pressure level (SPL). Level randomization was applied independently to the two complexes to prevent listeners from completing the task without

segregating the two complexes. Potentially, the listeners could detect a timbre change in *different* trials based on the mixture of the two complexes. However, the 20-dB rove made such a cue extremely unlikely. For example, simulations based on the output of a gammatone filter bank (Patterson *et al.*, 1995) indicated that for the stimuli tested here, a multichannel ideal observer failed to reach above-chance levels of performance when the PS was 40 dB, the largest value used in the current experiment.

All stimuli were generated digitally at a sampling frequency of 44 100 Hz on a personal computer (PC), which also controlled the experimental procedure and data collection through custom written MATLAB (The MathWorks, Inc., Natick, MA) software. The stimuli were presented to the listeners' two ears through the PC's 24-bit soundcard (Envy24 PCI audio controller, VIA Technologies, Inc., Taipei, Taiwan), a pair of programmable attenuators (PA4, Tucker-Davis Technologies, Inc., Alachua, FL), a headphone buffer (HB6, Tucker-Davis Technologies, Alachua, FL), and a pair of Sennheiser HD410 SL headphones (Old Lyme, CT). Each stimulus presentation was followed by a visual feedback indicating the correct response. The experiment was conducted in a double-walled, sound-attenuating booth.

2. Subjects

Five listeners with normal hearing (S1–S5), including the first author (S4), participated in this experiment. All listeners were between the ages of 18 and 30 and had audiometric thresholds at or better than 15 dB HL between 250 and 8000 Hz in both ears. Listeners were paid for their participation except for the author. The experiment was conducted in 2 h sessions. No more than one session was conducted for each listener on a single day. All listeners had participated in pilot projects of the current study, where they received extensive experiences in performing tasks very similar to the one used in the current experiments. Before data collection began, each listener practiced for at least 4 h before data collection started. These training sessions consisted of a subset of conditions from the experiment, which were combinations of 2 Δf_0 's (0.5 and 2 semitones), two onset delays (0 and 160 ms), and two Δ ITDs (0 and $\pm 400 \mu$ s). These eight conditions were tested in random order and repeated until the performance became consistently above chance (so that thresholds could be reliably estimated using a 2-down, 1-up procedure).

3. Procedure

Thresholds were estimated using a Same/Different procedure combined with 2-down, 1-up tracking algorithm, which estimated thresholds at the 70.7% point on the psychometric function (Levitt, 1971). Each track started at a PS of 30 dB, which was decreased after two consecutive correct responses and increased after a single incorrect response. The initial step size for these increments and decrements was 8 dB, which was reduced to 4 dB after the first two reversals. Each track terminated after a total of 10 reversals, and a threshold was estimated as the average of the PS values at the last six reversals.

A total of 50 conditions were tested (5 Δf_0 's \times 5 onset delays \times 2 Δ ITDs). Listeners S1, S4, and S5 ran 0-ITD conditions (including all five Δf_0 's and five onset delays) before starting the ± 400 -ITD conditions, whereas listeners S2 and S3 began with the ± 400 -ITD conditions. For each Δ ITD, the five onset delays were tested in random order. Within each delay, one track was run for each of the five Δf_0 's in random order. When a threshold estimate was obtained for each combination of Δf_0 and delay, the process was repeated three more times, using different randomizations. The resulting four threshold estimates in each condition were averaged to generate the reported thresholds. Then, the process was repeated for the remaining Δ ITD.

Thresholds were also estimated using a single complex. In this baseline condition, the spectral profile of the complex was either the same or different across the two intervals. Therefore, this condition measured the sensitivity to spectral-shape changes for an isolated stimulus. These thresholds provide estimates of the best performance (lowest thresholds) achievable by the listeners in the concurrent profile analysis task. In the case of perfect source segregation and no energetic masking, the concurrent profile analysis threshold should approach that obtained in the baseline condition. One baseline threshold using isolated stimuli was collected at the beginning of each experimental session on each subject. At least six thresholds were measured. The reported data were based on the average of the last four measurements.

B. Results

The concurrent profile analysis thresholds from the four listeners are plotted in Fig. 2 in separate rows. Results from 0- and ± 400 -ITD conditions (for Δ ITDs of 0 and 800 μ s, respectively) are plotted in the left and right columns. In each panel, thresholds are plotted as a function of Δf_0 , and various onset delays are indicated by different symbols.

Despite large individual differences (discussed later), general trends are present, as can be observed in the averaged thresholds across the five listeners shown in Fig. 3. Thresholds in the left-hand panels were generally higher than those in the right panels, indicating lower thresholds with ITD differences and higher thresholds when there were no ITD differences. Among the high thresholds in the left panels (0-ITD conditions), the highest thresholds were obtained in conditions with smallest Δf_0 's and shortest onset delays. Therefore, averaged across listeners, all three types of segregation cues improved thresholds in the concurrent profile analysis task.

A repeated-measure analysis of variance (ANOVA) was performed treating Δf_0 , onset asynchrony, and Δ ITD as the three within-subject factors. The ANOVA revealed significant main effects of onset asynchrony [$F(4,16) = 14.14$; $p < 0.001$] and Δf_0 [$F(4,16) = 8.16$; $p < 0.001$], and a marginal effect of Δ ITD [$F(1,4) = 7.89$; $p < 0.050$]. In general, these results agree with our expectation that introducing segregation cues, i.e., larger f_0 difference, ITD difference and onset delay, would yield lower thresholds. The sole significant interaction detected by the ANOVA was between Δf_0 and Δ ITD [$F(4,16) = 8.08$; $p < 0.001$]. This suggested that

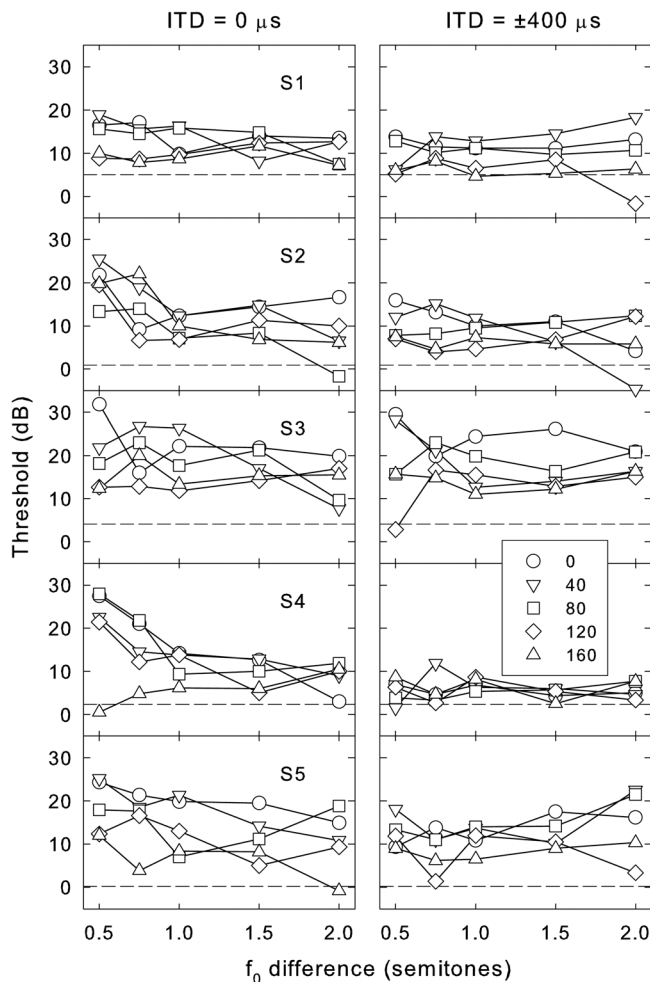


FIG. 2. Individual data from experiment I. Results for the 0 and ± 400 ITD conditions are shown in the left- and right-hand columns, respectively, and results for different listeners are in different rows. In each panel, thresholds are plotted as functions of f_0 difference, and various onset delays are indicated using different symbols. Dashed lines denote the isolated profile-analysis thresholds.

the effect of Δf_0 on threshold was smaller at larger Δ ITDs. Equivalently, the effect of Δ ITD was reduced when large Δf_0 's were present. It should be noted that this interaction may reflect a floor effect such that for large Δf_0 's or large Δ ITDs, thresholds cannot be further reduced.

Large individual differences are apparent in the pattern of thresholds (see Fig. 2). Potentially, different listeners may have adopted listening strategies that varied in terms of the relative importance among the three cue types, giving rise to different patterns of thresholds.

For example, listeners S3 and S4 showed very different patterns of thresholds. When Δ ITD and onset delay coexisted (see the third and fourth rows in Fig. 2), S3's thresholds tended to depend more on the amount of delay rather than Δ ITD, whereas S4's thresholds tended to rely more on Δ ITD. This observation might suggest that S3 weighted onset asynchrony more heavily in performing the task and S4 relied on Δ ITD information as the dominant segregation cue. To explore the possibility of different individual listening strategies, a stepwise multiple linear regression was performed on the individual thresholds for each listener with the independent variables being the amounts of Δf_0 , onset

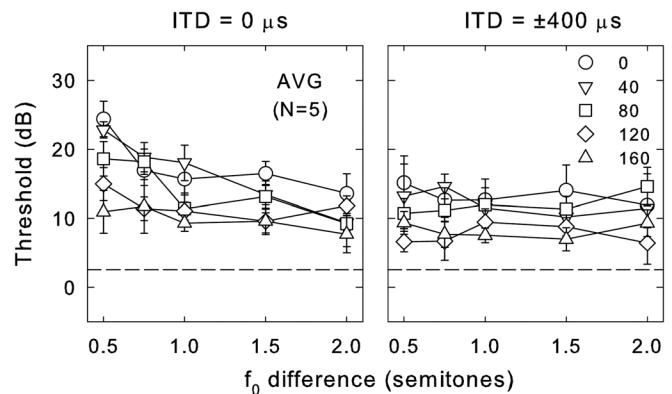


FIG. 3. Average thresholds across listeners from experiment I. The results are arranged in the same manner as in Fig. 3. Error bars indicate the standard errors of the means.

asynchrony, and Δ ITD, and the dependent variable being concurrent profile analysis threshold. The estimated regression coefficients and corresponding R^2 statistics for the three independent variables are listed in Table I.⁴ The significant coefficients in each fitted regression function are indicated using asterisks. As shown in the table, all estimated regression coefficients were negative, indicating that an increase in any of the three acoustic cues was associated with a decrease in threshold. Across the five listeners, the total variance accounted for by all three variables was similar (see the right-most column in Table I), ranging from 37% to 49%. However, the contributions from the three types of acoustic cues revealed different patterns for different listeners as illustrated by the R^2 values in parentheses. For example, although Δ ITD cues best accounted for the thresholds of S4, onset asynchrony was indicated as the most important cue for listener S3.

It is worth pointing out that when conducting regression analyses on the pooled data from all five listeners (250 thresholds: 5 listeners \times 5 Δf_0 's \times 5 onset delays \times 2 Δ ITDs), the "average" weighting strategy appeared to suggest that the three acoustic cues contributed to the performance more or less evenly (see R^2 values in the bottom row of Table I). This is contradictory to the observation that individual listeners tended to attend to one prominent cue type over the other two. For all listeners but S2, the patterns of thresholds suggest the dominance of one segregation cue over the others. Thus, one should be cautious in using pooled data to interpret the roles of the segregation cues in concurrent profile analysis.

III. EXPERIMENT II: CONCURRENT PROFILE ANALYSIS—PSYCHOMETRIC FUNCTIONS

In experiment II, psychometric functions were measured for the concurrent profile analysis task. These psychometric functions address the following questions: Whether energetic or informational masking dominates the performance in concurrent profile analysis, and how acoustic segregation cues affect the amount of masking.

Brungart (2001) measured speech intelligibility of target messages in the presence of competing masking messages.

TABLE I. Coefficients for the three types of segregation cues and R^2 statistics estimated using the stepwise linear regression for the threshold data from experiment I.^a

	ΔITD	Onset delay	Δf_0	Total R^2
S1	-0.007 (0.120)*	-0.043 (0.355)**	-0.882 (0.014)	0.488
S2	-0.010 (0.131)*	-0.024 (0.055)	-4.597 (0.187)**	0.373
S3	-0.000 (0.000)	-0.058 (0.335)**	-2.164 (0.042)	0.378
S4	-0.018 (0.316)**	-0.027 (0.056)*	-3.505 (0.086)*	0.458
S5	-0.005 (0.033)	-0.065 (0.399)**	-1.206 (0.012)	0.445
All	-0.008 (0.067)**	-0.043 (0.149)**	-2.471 (0.044)**	0.260

^aAsterisks indicate the significance of the variables in the regression analysis. * $p < 0.05$; ** $p < 0.001$. The values in parentheses indicate the proportion of variance accounted for (R^2) by each variable in the stepwise regression.

He found that listeners' errors were not random, as would be expected if energetic-masking dominated intelligibility. Rather the error pattern suggested that listeners were unable to distinguish the target from the masker, which led to informational masking. He also found that performance was better when different-sex target and masker voices were used compared to the same-sex speaker condition. Potentially, this was due to a Δf_0 -based source segregation. Because informational masking could play an important role in competing-speech recognition tasks, one might expect the involvement of informational masking in concurrent profile analysis experiments.

Psychometric functions have been utilized as a tool to reveal features of informational masking. For example, the occurrence of informational masking is often associated with shallow slopes of psychometric functions (see, e.g., Kidd *et al.*, 2003; Lutfi *et al.*, 2003; Durlach *et al.*, 2005), which reflect increased variability in decision-making process or internal noise. Moreover, for a number of trials, listeners might perceive multiple sources as a single auditory object, or they might not be able to selectively attending the target object (due to the similarity among the sources, attentional capacity limitations, etc). In such cases, the listeners would generate random responses unrelated to the detectability of the target. Such guessing leads to psychometric functions with upper asymptotes less than 1 (see, e.g., Green, 1995).

Figure 4 shows examples of psychometric functions from a computer simulated detection task. The three parameters that could affect the performance in this virtual task were the amount of energetic masking, the variance of the internal decision noise, and the proportion of trials where guessing occurred (see the figure caption for details about the simulation). The left-hand panel of Fig. 4 shows the effect of energetic masking on the psychometric function. More intense maskers shift the function horizontally toward higher signal levels (from dashed to solid curve in the panel) without significantly changing the slope and asymptote. In contrast, slope and asymptote changes are observed in the middle and right-hand panels of Fig. 4, illustrating two possible features of psychometric functions associated with informational masking. The slope of the psychometric function becomes shallower with increases in internal noise, whereas the asymptote falls when the proportion of guessing trials increases. Both of these effects could potentially lead

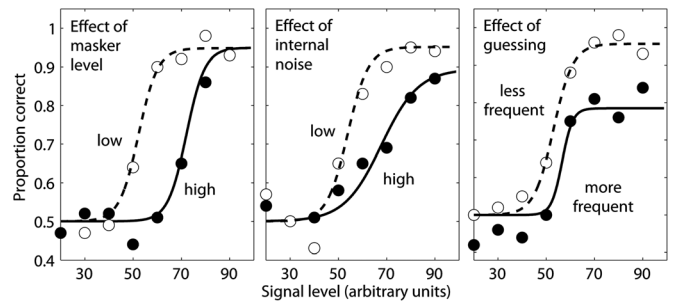


FIG. 4. Demonstration of the effects of energetic masking (left), internal decision variability (middle), and guessing (right) on the shape of the psychometric function. In each panel, circles are simulated proportion correct data at various signal levels in a virtual detection task. Each data point is based on 100 simulated trials. Within each trial, either a random response was generated based on the probability of confusion, p_{guess} , or a detection was reported when $10 \log(10^{L_s/10} + 10^{L_m/10}) + \epsilon_1 > L_m + \epsilon_2$, where L_s and L_m are the signal and masker levels respectively, ϵ_1 and ϵ_2 are two independent random values drawn from a distribution of internal noise with a mean of zero and a standard deviation of σ_0 . The open and filled symbols in the three panels, from left to right, are for two different values of L_m , σ_0 , and p_{guess} , respectively. The two sets of simulated data are fitted with two psychometric functions (dashed and solid curves for open and filled data, respectively) using the same procedure described in Sec. III A.

to an increase in threshold without increasing masker intensity. Therefore, studying these two features (slope and asymptote) of the psychometric function has the potential to reveal possible origins of informational masking.

The psychometric functions estimated in the present experiment provide an opportunity to study the role of informational masking in concurrent profile analysis, and the systematic effects of segregation cues on informational masking.

A. Methods

Five normal-hearing listeners (S2–S6) participated in this experiment, four of whom participated in experiment I. Due to the limited availability of listener S1, a new listener S6 was recruited. This new participant was 19 years of age and had audiometric thresholds at or better than 15 dB HL between 250 and 8000 Hz in both ears. Similar to other listeners, listener S6 was also an experienced listener for profile analysis tasks. Before data collection began, this listener received training on the concurrent profile analysis task in the same way that the other listeners were trained before experiment I. Experiment II was conducted in 2 h sessions. No more than one session was conducted for each listener on a single day.

Listeners' responses as to whether a *same* or *different* trial was presented (see Fig. 1) were recorded in blocks of 60 trials, which contained ten trials at each of six fixed PS values (−10, 0, 10, 20, 30, and 40 dB). Before a block started, the sequence of the 60 trials in a block was determined randomly. In all regards the stimulus generation and presentation methods were as in experiment I. The Δf_0 's tested were 0.5 and 2 semitones, the onset delays tested were 0 and 160 ms, and the ΔITD s tested were 0 and 800 μs , forming a total of eight conditions. Eight experimental blocks corresponding to the eight conditions were run in random order, and then the process was repeated five times, each time with a different sequence.

Therefore, for each of the eight conditions, 50 responses were collected at each of the six PS values, from which the proportion of the correct responses was calculated. Proportion correct at the same six PS values (based on 50 responses at each PS value collected in five blocks) was also measured for an isolated complex to provide a baseline psychometric function when no masking was present.

The psychometric functions were characterized by fitting logistic functions to the proportion correct data using a maximum-likelihood criterion as suggested by [Wichmann and Hill \(2001a\)](#) and [Lutfi et al. \(2003\)](#). The MATLAB software package PSIGNFIT ([Wichmann and Hill, 2001a,b](#)) was used. The form of the fitted psychometric function was

$$\hat{P}_c = 0.5 + (0.5 - \lambda) \left(1 + e^{-(PS-\alpha)/\beta} \right)^{-1}, \quad (3)$$

where \hat{P}_c is the estimated proportion correct, λ describes the distance from the upper asymptote of the function to 100% correct, α denotes the horizontal position of the function, and β gives the slope of the function.⁵ Note that the slope decreases with larger β values.

As an example, this fitting procedure was applied to the simulated data shown in Fig. 4, and the fitted psychometric functions are shown as the solid and dashed curves. As mentioned previously, the simulated energetic masking shifts the psychometric function horizontally (left-hand panel). This corresponds to increasing α values with increasing amounts of the simulated energetic masking. On the other hand, when informational masking occurs, in this simulation, the psychometric functions exhibit shallower slopes (middle panel) and lower asymptotes (right-hand panel), which correspond to larger β and λ values, respectively.

In the present experiment, the major focus is the role of informational masking in the concurrent profile analysis task. This was examined by asking whether manipulating segregation cues led to systematic changes to the values of β and λ . Further, if informational masking limits performance in the concurrent profile analysis task, the effects of the three types of acoustic cues on the values of β and λ might reveal how these cues contribute to the release from informational masking. For example, if increases in the magnitude of segregation cues yield reduction in β , it would imply that the segregation cues improve performance by reducing the magnitude of internal noise. On the other hand, a correlation between acoustic cues and the value of λ would suggest that these cues reduce the proportion of random guessing.

In order to reveal such relations, stepwise multiple linear regressions were performed on β and λ . Two regressions were run for each listener. The independent variables were the sizes of Δf_0 , onset asynchrony, and ΔITD , and the dependent variable was the estimates of β or λ . The regressions provided estimates of the portion of the total variance in β or λ accounted for by each of the segregation cues for each listener. The resulting R^2 values indicated the relative contributions of the acoustic cues to changes in the slope and asymptote of the psychometric function.

Because β and λ were estimated from the fitted psychometric functions, the dependent variable of the stepwise regression is not known with certainty. This, in turn, leads to

nontraditional distributions of R^2 values. To provide an estimate of the distribution of R^2 values, a bootstrap procedure was implemented. For each experimental condition, β and λ values were drawn 1000 times from the distributions estimated when the psychometric functions were fitted.⁶ For each of the 1000 replicates, a stepwise regression was conducted on the eight resampled β values (corresponding to 2 $\Delta ITDs \times 2$ delays $\times 2$ Δf_0 's), and a separate regression was performed on the resampled λ values.⁷ The resulting list of 1000 R^2 values revealed the distribution of the R^2 statistics.

This regression analysis was also applied to the data set as a whole. In this case, during each resampling-and-regression process, 40 resampled values were drawn from all 40 distributions (8 conditions \times 5 listeners) of either β or λ . The 40 resampled values were then entered into a stepwise linear regression in order to calculate a R^2 value in each of the eight conditions. The distributions of the R^2 values were obtained after repeating the process 1000 times.

B. Results

As an example, Fig. 5 shows the proportion correct as functions of PS in each of the eight conditions (filled symbols) as well as the fitted psychometric functions (solid and dash-dotted curves) for one of the listeners (S6). The two $\Delta ITDs$ are arranged in columns, the two onset delays are arranged in rows, and the filled squares and diamonds denote Δf_0 's of 0.5 and 2 semitones, respectively. Also plotted in Fig. 5 are proportion correct and fitted psychometric functions from the isolated profile-analysis task (as open circles and dashed curves, respectively). Compared to this baseline condition, the proportion correct was lower in the concurrent conditions in general, which led to shallower slopes or lower asymptotes in the estimated psychometric functions. As discussed earlier, shallow psychometric functions with low asymptotes suggest the involvement of informational masking in the concurrent profile analysis task. One distinct feature that can be observed from Fig. 5 was that the asymptotes of the psychometric functions, or the best possible performance, were very low in some conditions. The maximum proportion correct was as low as 0.6 in the condition with smallest ΔITD , delay, and Δf_0 values. This suggests that the listener responded by guessing on more than half of the trials, indicating the extreme difficulty of the task. Comparing among conditions for this listener (S6), one consistent trend was that onset asynchrony seemed to affect the asymptotes of the psychometric functions. The maximum proportion correct increased about 10 percentage points when the onset delay was increased from 0 (top panels) to 160 (bottom panels) ms. The value of Δf_0 also appeared to affect the shapes of the psychometric functions. However, given the very low percentage correct at the Δf_0 of 0.5 semitones, poor psychometric-function fits were often obtained (squares and dash-dotted curves in the left-hand panels), making it difficult to assess the effect of Δf_0 .

The psychometric functions fitted to each listener's data were analyzed quantitatively using the values for the parameters α , β , and λ [see Eq. (3)]. By its definition, the parameter α was the profile strength at the center of the psychometric

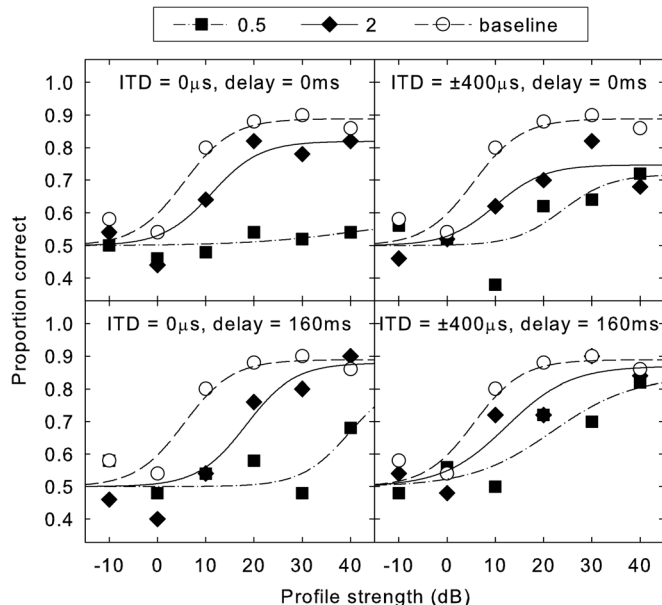


FIG. 5. The proportion correct (symbols) and estimated psychometric functions (curves) for listener S6 in experiment II. The filled symbols are based on the eight concurrent profile analysis conditions differing in ΔITDs (left- and right-hand panels), onset delays (top and bottom panels), and Δf_0 's (as different symbols). The open symbols are for the isolated profile-analysis condition, and are the same across the four panels.

function's dynamic range (75% point if $\lambda = 0$). Therefore, it could be considered as a threshold measure. Indeed, for the four listeners who participated in both experiments, the estimated α values were very closely to the thresholds collected in experiment I using 2-down, 1-up tracking procedure (70.7% point on the psychometric function). This indicates good consistency across the two experiments.

Table II shows the values of β and λ estimated from all listeners in all conditions. The standard errors of the estimates are indicated in parentheses. A dash (-) in Table II indicates poor psychometric function fits (defined as fits yielding values of α larger than 40 dB). Large individual differences were observed in the estimated β and λ values, which was expected given the results of experiment I. Therefore, the effects of the segregation cues, visually observed in

Fig. 5 for listener S6, were not consistent across listeners. For example, onset delay affected λ for S6 (larger the delay smaller the λ) but showed no notable effect on λ for S4.

To investigate the relative contributions of ΔITD , onset asynchrony, and Δf_0 on β and λ for each listener, in a bootstrap procedure, 1000 values of β and λ were drawn from their estimated distributions (Table II) and then used as the dependent variables in two separate stepwise multiple linear regression analyses (see Sec. III A for computational details). The distributions of the resulting R^2 statistics are plotted in Figs. 6 and 7 for β and λ , respectively. Results for the five individual listeners and results derived from the β and λ estimates using the data from all listeners are plotted in separate panels. The R^2 values illustrate the relative strength of each segregation cue, accounting for the total variance in the resampled β or λ . The cue that possesses a larger R^2 has a relatively dominant role in accounting for either the slope (β) or asymptote (λ) of the psychometric function.

The estimated R^2 values were distributed over a wide range. However, information can be gained by studying the shapes of these R^2 distributions, whose 25th, 50th, and 75th percentiles are indicated by the box plot in each panel. For β (Fig. 6), R^2 estimates from each of the three cue types were similar. For listeners S2 and S6, the slope of the psychometric function was more closely associated with ΔITD , whereas for S3 and S4, Δf_0 exhibited a somewhat higher correspondence to the slope. The R^2 pattern estimated based on all listeners' psychometric functions (the bottom right panel in Fig. 6) also showed a slight advantage of Δf_0 over the other cues in accounting for the β and the slope of the psychometric function.

For λ (Fig. 7), the asymptote of the psychometric function seemed to be dominated by onset asynchrony between the two concurrent complexes for all listeners except listener S4, whose psychometric functions approached 100% correct (i.e., very small λ estimates) in most of the conditions. The dominance of onset asynchrony was also observed when pooling all listeners' data in the regression analysis (the bottom right-hand panel in Fig. 7). Therefore, introducing onset asynchrony tends to reduce the proportion of responses that

TABLE II. The β and λ parameters estimated from the psychometric functions in experiment II.^a

ΔITD (μs)	Delay (ms)	Δf_0 (semitones)	S2		S3		S4		S5		S6	
			β	λ	β	λ	β	λ	β	λ	β	λ
0	0	0.5	-	-	-	-	20.5 (19.9)	0.00 (0.00)	-	-	-	-
0	0	2	11.5 (10.4)	0.19 (0.13)	0.73 (1.51)	0.32 (0.04)	0.7 (0.3)	0.10 (0.02)	1.8 (4.8)	0.23 (0.04)	0.3 (0.2)	0.19 (0.03)
0	160	0.5	-	-	4.25 (5.42)	0.19 (0.04)	4.7 (2.7)	0.08 (0.03)	3.2 (3.2)	0.15 (0.04)	-	-
0	160	2	6.0 (7.5)	0.20 (0.04)	7.14 (9.10)	0.22 (0.05)	7.6 (2.7)	0.03 (0.03)	4.9 (2.4)	0.05 (0.03)	3.7 (3.7)	0.14 (0.04)
400	0	0.5	3.9 (6.5)	0.24 (0.05)	9.89 (11.18)	0.22 (0.14)	8.0 (2.9)	0.03 (0.03)	-	-	1.0 (5.9)	0.32 (0.05)
400	0	2	0.9 (1.3)	0.19 (0.04)	0.99 (6.36)	0.29 (0.04)	1.3 (1.7)	0.02 (0.01)	4.6 (6.7)	0.22 (0.05)	4.2 (7.3)	0.26 (0.04)
400	160	0.5	0.8 (0.6)	0.16 (0.03)	5.78 (6.00)	0.18 (0.06)	3.3 (1.7)	0.03 (0.02)	5.3 (3.0)	0.08 (0.04)	8.4 (7.4)	0.16 (0.12)
400	160	2	8.6 (8.2)	0.14 (0.04)	2.22 (3.51)	0.16 (0.03)	3.9 (1.9)	0.03 (0.02)	0.8 (1.0)	0.11 (0.03)	6.6 (4.5)	0.13 (0.05)
Baseline			3.8 (5.1)	0.20 (0.04)	4.41 (3.59)	0.18 (0.03)	0.9 (1.6)	0.03 (0.01)	0.0 (0.0)	0.00 (0.00)	3.1 (3.0)	0.12 (0.03)

^aThe values in parentheses are the standard error of the estimates. A dash (-) indicates poor psychometric-function fit (defined as fits yielding values of α larger than 40 dB). These conditions were excluded from further statistical analysis. This occurred in four conditions from three listeners. These conditions also showed very low maximum performance, typically less than 75% correct.

result from guessing, potentially by enhancing source segregation.

Comparing across experiments I and II, the influences of acoustic cues on thresholds seemed to be consistent with their effects on the λ estimates. For examples, the regression results in Table I suggested that listeners S3 and S5 adopted onset asynchrony as the dominant cue for segregation. For these two listeners, onset asynchrony also best accounted for the λ estimates the best (Fig. 7). Moreover, Δ ITD was the dominant cue for listener S4 in experiment I, this cue also exhibited the strongest association with the λ estimates. Therefore, in the small sample of four listeners that participated in both experiments, three listeners showed consistent effects of the dominant cue on the estimated threshold and λ values. For these three listeners, the introduction of the dominant cue seemed to improve the threshold by alleviating listeners' confusion, thereby reducing the proportion of random responses.

IV. GENERAL DISCUSSION

A. Agreement with previous studies using double vowels

The present study developed a new psychophysical paradigm, concurrent profile analysis, to access sensitivity to spectral profiles of concurrent sound stimuli as an alternative to the traditional double-vowel paradigm. With this newly developed method, the effect of segregation cues on spectral

processing can be investigated using nonspeech stimuli, hence closed-set stimulus designs are not required. Further, through the concurrent profile analysis task, psychometric functions can be measured, which provide more detailed information regarding the involvement of informational masking and possible causes of individual differences in performance.

There is a good agreement between the findings of the current experiments and double-vowel studies. Segregation cues, such as Δf_0 , onset asynchrony, and Δ ITD, improve the performance in both double-vowel identification and the detection of spectral-profile changes. For Δf_0 , several studies have shown that double-vowel identification performance increases with increasing Δf_0 between 0 and 1 semitones, and asymptotes for larger Δf_0 's (see, e.g., Assmann and Summerfield, 1990; Culling and Darwin, 1993). Similarly, most of the listeners in experiment I showed Δf_0 -based improvement in thresholds, especially when Δf_0 was the only available segregation cue (see the circles in the left-hand panels of Fig. 2). Moreover, the influence of Δf_0 on threshold is most evident for Δf_0 below 1 semitone, in agreement with the pattern of Δf_0 -based improvement in double-vowel experiments.

For the effect of onset asynchrony, double-vowel experiments typically show an increase of identification scores with increasing onset delays up to 200 ms before reaching an asymptote (see, e.g., Lentz and Marsh, 2006). A similar benefit from onset asynchrony was observed in experiment I of the present study (see the third column of Table I). Lentz and Marsh (2006) investigated the effect of Δf_0 and onset asynchrony simultaneously using Δf_0 's of 0 and 4 semitones and onset delays ranged from 0 to 300 ms. They found that the benefit on vowel identification of onset asynchrony did not depend on Δf_0 . In experiment I, the interaction between Δf_0 and onset asynchrony was not significant, in agreement with the results of Lentz and Marsh (2006).

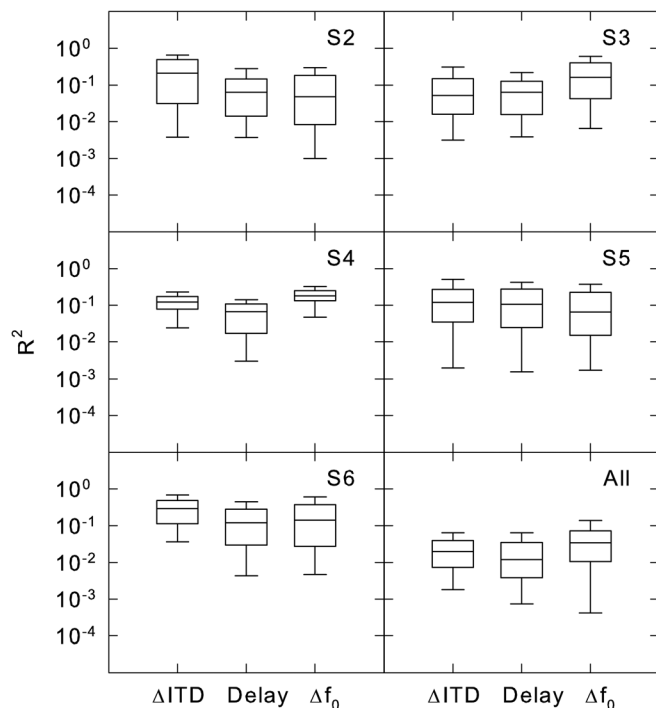


FIG. 6. The distributions of R^2 statistics for β estimated using a bootstrap procedure in experiment II. The bootstrap procedure computed the R^2 distributions by repeating a resampling and regression process 1000 times. Within each repetition, a stepwise linear regression was conducted on the resampled β data. Results for individual listeners and results for the data from all listeners are shown in separate panels. Within each panel, the horizontal line within each box indicates the median of the R^2 distribution. The boundaries of the boxes indicate the 25th and 75th percentiles, whereas the whiskers below and above the boxes indicate the 10th and 90th percentiles.

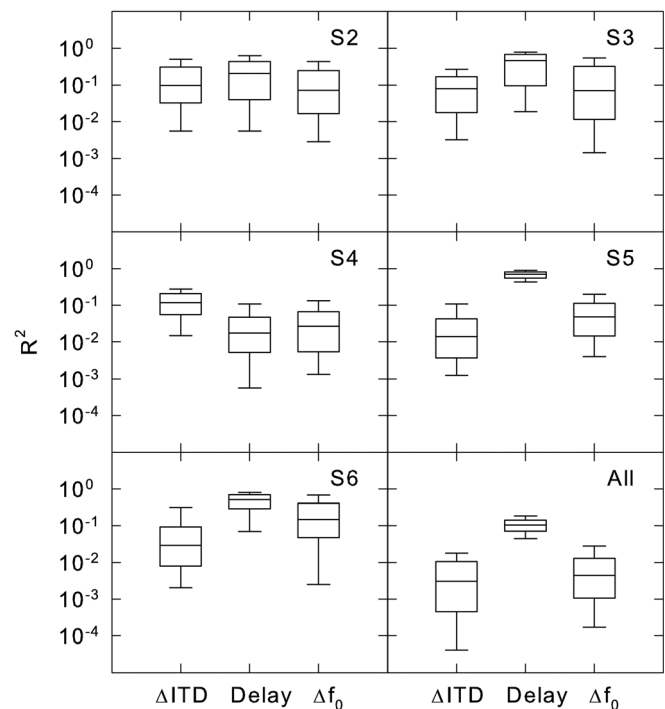


FIG. 7. Same as Fig. 6, but for the resampled λ data in experiment II.

Regarding ITD differences, it has been argued that Δ ITD is a weaker segregation cue compared to Δf_0 and onset asynchrony, resulting in less benefit for double-vowel identification (e.g., Shackleton and Meddis, 1992). In experiment I, this weaker role of Δ ITD is reflected in the fact that the effect of Δ ITD is less prominent than the two other cue types. Among the five listeners, Δ ITD best accounted for the pattern of thresholds for only one of the listeners (see R^2 values for listener S4 in Table I). A few studies have investigated the effect of Δ ITD and Δf_0 simultaneously. For example, Shackleton *et al.* (1994) used Δf_0 's of 0 and 1 semitones and ITDs from 0 to $\pm 400\mu\text{s}$ in their double-vowel experiment and found that a Δf_0 was required for listeners to receive benefit from Δ ITD.⁸ In experiment I, we found a significant interaction between Δ ITD and Δf_0 , which agreed with the data of Shackleton *et al.* (1994).

In summary, results from experiment I are in agreement with previous double-vowel studies in terms of the benefit from the three types of segregation cues and their interactions. However, a smaller number of listeners participated in the current experiments and the individual differences were larger.

B. Limitations to concurrent profile analysis paradigm

Through two experiments, promising results were obtained in the current study, suggesting the potential of using the concurrent profile analysis paradigm as a tool to study concurrent spectral processing. However, one should be aware that a number of issues could potentially confound the interpretations of the concurrent profile analysis data.

The threshold data obtained in experiment I and the psychometric-function data obtained in experiment II were not always consistent with each other. It was shown in experiment II that the upper asymptote of the psychometric function could be very low in some conditions. The best performance that a listener could achieve was not always above 70.7% correct. How, then, could thresholds be estimated at 70.7% correct in experiment I? Moreover, for three of the five listeners that participated in experiment II, the asymptote in the psychometric function was fairly low, even in the baseline condition (see λ values in the bottom row of Table II). It is likely that these observed discrepancies reflect the fact that the subjects' task was more difficult in experiment II than experiment I. In experiment I, the profile strength applied on the current trial was in the vicinity of the profile strength on the last trial; whereas in experiment II, the profile strength on each trial was randomly drawn across a large range (from -10 to 40dB). It is possible, therefore, that a portion of apparent informational masking was a result of task demands and not necessarily caused by the introduction of a competing sound source. Further, in both experiments, various randomizations implemented in the experimental design prevented listeners from having *a priori* knowledge of the values of f_0 's and sound levels of the two complexes. This uncertainty, which prevented the use of unwanted acoustic cues, had the potential of introducing extra informational masking that was unrelated to the failure of segregation.

C. Implications

Psychometric functions for concurrent profile analysis measured in experiment II suggested that informational masking is an important source of masking for the task. It could be hypothesized based on the results of experiment II that the presence of competing sound sources increases both internal noise and confusion among the sources leading to shallow psychometric functions with low asymptotes. This means that in a competing-source situation, enhancing profile strength might have little effect in improving spectral sensitivity. On the other hand, enhancing segregation cues might alleviate the amount of informational masking and promote better performance.

One of the signal-processing schemes developed in recent years for improving speech recognition in noise through assistive listening devices is spectral-contrast enhancement, or spectral enhancement. Motivated by the fact that hearing-impaired listeners need higher peak-to-trough ratios in the spectrum to correctly identify steady-state vowels (Leek *et al.*, 1987), spectral-enhancement algorithms selectively amplify formants and attenuate spectral troughs. Several studies have been carried out to explore the efficacy of spectral enhancement (see, e.g., Baer *et al.*, 1993; Franck *et al.*, 1999; Lyzenga *et al.*, 2002). Significant benefits from spectral enhancement were not reliably found in these studies (see, e.g., DiGiovanni *et al.*, 2005). The results of experiment II provide insights into these mixed findings. It is possible that the benefit of spectral enhancement schemes could be undermined owing to the large amount of informational masking, especially in situations where the target speech is embedded in spectrotemporally complex maskers or competing speech messages. A large amount of informational masking is usually associated with very shallow psychometric functions and low asymptotes, such as those observed in experiment II. That is, listeners might be able to hear the individual spectral peaks clearly enough (given the spectral enhancement), however, poor segregation of the target and masker or failures of selectively attending the target might be the limiting factor to the overall performance. The concurrent profile analysis task developed in the current study provides a viable tool to study this possibility in the future.

V. SUMMARY

A concurrent profile analysis task was developed to provide a measure of sensitivity to spectral changes in competing sound sources. Experiment I demonstrated that larger values of Δf_0 , onset asynchrony, and Δ ITD between the two concurrent complexes led to lower thresholds. When segregation cues were jointly present, the dominant cue differed across listeners. These different strategies in using the cues gave rise to different patterns of results among the listeners. In experiment II, it was found that all three types of segregation cues contributed to changes in the slope and asymptote of the psychometric function, with increases in the magnitude of the cues leading to steeper slopes and high asymptotes, suggesting the involvement of informational masking. Individual differences were also evident in this experiment. Despite large individual differences, one consistent trend across most of the listeners (four out of five listeners) was

that onset asynchrony plays an essential role in determining the asymptote in the psychometric function. This is in fair agreement with the results of experiment I, which indicated that onset delay dominated the concurrent profile analysis threshold for three out of five listeners.

ACKNOWLEDGMENTS

This work was supported by NIH Grant Nos. R01 DC02012 and R21 DC010058. The authors would like to thank Eva Maria Carreira and Andrew Silva for their assistance in data collection.

¹Note that this statement is not necessarily true without avoiding the involvement of cues that allow differentiating *different* from *same* trials based on the mixture of the two complexes. In the current experiment, this was done by implementing level and fundamental frequency randomization (see Sec. II A for details).

²ITD is the major cue for lateralization at low frequencies, whereas interaural level difference (ILD) is the dominant cue for lateralization at high frequencies (e.g., Sandel *et al.*, 1955). In the current experiment, broadband stimuli were used and the ILD was fixed at 0 dB, therefore a conflict might arise between the ITD and ILD cues. However, in a double-vowel experiment, Shackleton *et al.* (1994) showed that applying or removing ILD cues, the benefit from Δ ITD on vowel identification was not significantly affected. Therefore, we expect the influence of the conflicting ITD and ILD cues to be limited in our experiment.

³The time shifts in the current experiment for ITD generation was achieved via manipulating the phase spectra of the stimuli in the frequency domain. A Fourier transform was performed on a stimulus after it had been zero-added before its temporal onset and after its offset, 1 ms was added to each end. The phase spectrum was then shifted according to a linear function of frequency (Hz), the slope of which was the intended time delay (s).

⁴In the regression analysis, the variables were in units of semitones, milliseconds, and microseconds for the Δf_0 , onset asynchrony, and Δ ITD cues, respectively.

⁵ λ can also influence the slope of psychometric function, but to a lesser degree than β (Lutfi *et al.*, 2003).

⁶The PSIGNIFT software package (Wichmann and Hill, 2001a,b) used in the current study to estimate the psychometric functions provides not only estimates of β and λ , but also their confidence intervals obtained through a bootstrap procedure. For the purpose of resampling, the β and λ were assumed to have normal distributions, truncated at the parameter limits ($\beta > 0$ and $0 \leq \lambda \leq 0.5$). The standard deviation of each of these normal distributions was set to be half of the distance between the two 68% confidence limits provided by the PSIGNIFT software.

⁷Some conditions were removed from the analysis for some listeners owing to very poor psychometric-function fits where the estimated α was greater than 40 dB, the largest PS value tested in the experiment.

⁸In an earlier study (Shackleton and Meddis, 1992), this interaction between Δ ITD and Δf_0 was not found to be significant. A discussion of these inconsistent results can be found in Shackleton *et al.* (1994).

Assmann, P. F., and Summerfield, Q. (1989). "Modeling the perception of concurrent vowels: Vowels with the same fundamental frequency," *J. Acoust. Soc. Am.* **85**, 327–338.

Assmann, P. F., and Summerfield, Q. (1990). "Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **88**, 680–697.

Baer, T., Moore, B. C., and Gatehouse, S. (1993). "Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality, and response times," *J. Rehabil. Res. Dev.* **30**, 49–72.

Bernstein, J. G. W., and Oxenham, A. J. (2008). "Harmonic segregation through mistuning can improve fundamental frequency discrimination," *J. Acoust. Soc. Am.* **124**, 1653–1667.

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge, MA).

Brungart, D. S. (2001). "Informational and energetic masking effects in the perception of two simultaneous talkers," *J. Acoust. Soc. Am.* **109**, 1101–1109.

Carlyon, R. P. (1996a). "Encoding the fundamental frequency of a complex tone in the presence of a spectrally overlapping masker," *J. Acoust. Soc. Am.* **99**, 517–524.

Carlyon, R. P. (1996b). "Masker asynchrony impairs the fundamental-frequency discrimination of unresolved harmonics," *J. Acoust. Soc. Am.* **99**, 525–533.

Carlyon, R. P. (1997). "The effects of two temporal cues on pitch judgments," *J. Acoust. Soc. Am.* **102**, 1097–1105.

Culling, J. F., and Darwin, C. J. (1993). "Perceptual separation of simultaneous vowels: Within and across-formant grouping by F_0 ," *J. Acoust. Soc. Am.* **93**, 3454–3467.

Darwin, C. J. (1984). "Perceiving vowels in the presence of another sound: Constraints on formant perception," *J. Acoust. Soc. Am.* **76**, 1636–1647.

Darwin, C. J., and Hukin, R. W. (1997). "Perceptual segregation of a harmonic from a vowel by interaural time difference and frequency proximity," *J. Acoust. Soc. Am.* **102**, 2316–2324.

de Cheveigné, A., Kawahara, H., Tsuzaki, M., and Aikawa, K. (1997). "Concurrent vowel identification. I. Effects of relative amplitude and F_0 difference," *J. Acoust. Soc. Am.* **101**, 2839–2847.

DiGiovanni, J. J., Nelson, P. B., and Schlauch, R. S. (2005). "A psychophysical evaluation of spectral enhancement," *J. Speech Lang. Hear. Res.* **48**, 1121–1135.

Durlach, N. I., Mason, C. R., Gallun, F. J., Shinn-Cunningham, B., Colburn, H. S., and Kidd, G. (2005). "Informational masking for simultaneous non-speech stimuli: Psychometric functions for fixed and randomly mixed maskers," *J. Acoust. Soc. Am.* **118**, 2482–2497.

Durlach, N. I., Mason, C. R., Kidd, G., Arbogast, T. L., Colburn, H. S., and Shinn-Cunningham, B. (2003). "Note on informational masking," *J. Acoust. Soc. Am.* **113**, 2984–2987.

Franck, B. A., van Kreveld-Bos, C. S., Dreschler, W. A., and Verschuure, H. (1999). "Evaluation of spectral enhancement in hearing aids, combined with phonemic compression," *J. Acoust. Soc. Am.* **106**, 1452–1464.

Green, D. M. (1983). "Profile analysis: A different view of auditory intensity discrimination," *Am. Psychol.* **38**, 133–142.

Green, D. M. (1995). "Maximum-likelihood procedures and the inattentive observer," *J. Acoust. Soc. Am.* **97**, 3749–3760.

Green, D. M., and Kidd, G. (1983). "Further studies of auditory profile analysis," *J. Acoust. Soc. Am.* **73**, 1260–1265.

Green, D. M., Onsan, Z. A., and Forrest, T. G. (1987). "Frequency effects in profile analysis and detecting complex spectral changes," *J. Acoust. Soc. Am.* **81**, 692–699.

Hedrick, M. S., and Madix, S. G. (2009). "Effect of vowel identity and onset asynchrony on concurrent vowel identification," *J. Speech Lang. Hear. Res.* **52**, 696–705.

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.

Hukin, R. W., and Darwin, C. J. (1995). "Comparison of the effect of onset asynchrony on auditory grouping in pitch matching and vowel identification," *Percept. Psychophys.* **57**, 191–196.

Kidd, G., Mason, C. R., and Richards, V. M. (2003). "Multiple bursts, multiple looks, and stream coherence in the release from informational masking," *J. Acoust. Soc. Am.* **114**, 2835–2845.

Leek, M. R., Dorman, M. F., and Summerfield, Q. (1987). "Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **81**, 148–154.

Lentz, J. J., and Marsh, S. L. (2006). "The effect of hearing loss on identification of asynchronous double vowels," *J. Speech Lang. Hear. Res.* **49**, 1354–1367.

Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**(Suppl. 2), 467–477.

Lutfi, R. A., Kistler, D. J., Callahan, M. R., and Wightman, F. L. (2003). "Psychometric functions for informational masking," *J. Acoust. Soc. Am.* **114**, 3273–3282.

Lyzenga, J., Festen, J. M., and Houtgast, T. (2002). "A speech enhancement scheme incorporating spectral expansion evaluated with simulated loss of frequency selectivity," *J. Acoust. Soc. Am.* **112**, 1145–1157.

Meddis, R., and Hewitt, M. J. (1992). "Modeling the identification of concurrent vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **91**, 233–245.

Micheyl, C., Bernstein, J. G. W., and Oxenham, A. J. (2006). "Detection and f_0 discrimination of harmonic complex tones in the presence of competing tones or noise," *J. Acoust. Soc. Am.* **120**, 1493–1505.

- Micheyl, C., and Oxenham, A. J. (2010). "Pitch, harmonicity and concurrent sound segregation: Psychoacoustical and neurophysiological findings," *Hear. Res.* **266**, 36–51.
- Patterson, R. D., Allerhand, M. H., and Giguère, C. (1995). "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," *J. Acoust. Soc. Am.* **98**, 1890–1894.
- Qian, J., and Richards, V. M. (2010). "The effect of onset asynchrony on relative weights in profile analysis," *J. Acoust. Soc. Am.* **127**, 2461–2465.
- Rasch, R. A. (1978). "Perception of simultaneous notes such as in polyphonic music," *Acustica* **40**, 21–33.
- Sandel, T. T., Teas, D. C., Feddersen, W. E., and Jeffress, L. A. (1955). "Localization of sound from single and paired sources," *J. Acoust. Soc. Am.* **27**, 842–852.
- Shackleton, T. M., and Meddis, R. (1992). "The role of interaural time difference and fundamental frequency difference in the identification of concurrent vowel pairs," *J. Acoust. Soc. Am.* **91**, 3579–3581.
- Shackleton, T. M., Meddis, R., and Hewitt, M. J. (1994). "The role of binaural and fundamental-frequency difference cues in the identification of concurrently presented vowels," *Q. J. Exp. Psychol.* **47**, 545–563.
- Wichmann, F. A., and Hill, N. J. (2001a). "The psychometric function: I. Fitting, sampling, and goodness of fit," *Percept. Psychophys.* **63**, 1293–1313.
- Wichmann, F. A., and Hill, N. J. (2001b). "The psychometric function: II. Bootstrap-based confidence intervals and sampling," *Percept. Psychophys.* **63**, 1314–1329.
- Zwicker, U. T. (1984). "Auditory recognition of diotic and dichotic vowel pairs," *Speech Commun.* **3**, 265–277.