

Borrower: IUI

Lending String:

*CGU,OBE,WSU,YUS,LRU,HLS,TFW,ZCU,RCE,
VRC,QE2,UTO,SLY
In Process: 20180406

Patron:

Reference #: 10811874

Journal Title: Mathematical models of perception
and cognition /

Volume: 2

Issue:

Month/Year: 2016

Pages:

Article Author: Joseph W. Houpt and Leslie M.
Blaha

Article Title: Processing Characteristics of
Monaural Tone Detection: A Reaction Time
Perspective on a Classic Psychoacoustic Problem

Imprint: New York : Routledge, Taylor & Francis
Group, 2016.

Notes:

ILL Number: 187068859



ILLiad TN: 2230419

University of Chicago Interlibrary Loan
OCLC CAT / RICHCO / DOCUMENT

f

Call #: BF39 .M38 2016 c.1 v.2

Location: JRL / Gen

OCLC

CIC

Charge

Maxcost: 45.00IFM

Category: Exempt
Billing Notes

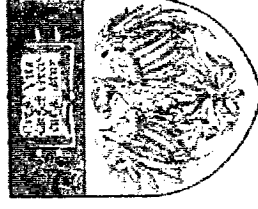
Shipping Address:

DDS, Wells Library, E-065
Indiana University Libraries
1320 E. 10th Street
Bloomington, Indiana 47405
United States

Fax: 812-855-8229

Ariel: ODYSSEY

Email: libill@indiana.edu



ODYSSEY REQUEST

SCAN for PLATES (Pattern 2)

Email / Ariel

Mail

Call Number: 416/2018 12:56:26 PM

Need By

Not Wanted After: 05/11/2018

Received Via:

SENT

Notice: This material may be protected by copyright
law (Title 17 US Code) Yes

Please report all Ariel transmission
problems within 48 hours of receipt

- van Wassenhove, V., Grant, K., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences, U. S.A.*, *102*, 1181–1186.
- Wenger, M. J., & Townsend, J. T. (2006). On the costs and benefits of faces and words: Process characteristics of feature search in highly meaningful stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 755–779.
- Zekveld, A. A., George, E. L., Houtgast, T., & Kramer, S. E. (2013). Cognitive abilities relate to self-reported hearing disability. *Journal of Speech, Language, & Hearing Research*, *56*, 1364–1372.

7 Processing Characteristics of Monaural Tone Detection

A Reaction-Time Perspective on a Classic Psychoacoustic Problem

Jennifer J. Lentz, Yuan He, Joseph W. Houpt, Julia M. DeLong, and James T. Townsend

7.1 Introduction

One of the most fundamental characteristics of the auditory system is that of frequency analysis. At the level of the cochlea, the basilar membrane can be likened to a bank of filters that separates the various frequencies of complex sounds akin to Fourier analysis. In this distributed system, high-frequency sounds are represented at one end of the cochlea (the base, near the stapes), whereas low-frequency sounds are represented at the apical end.

This tonotopic organization is maintained throughout all levels of the auditory system in which different neural populations are tuned to different frequencies. Notably, perception is that of unified complex sounds. Consequently, one of the fundamental research endeavors in psychoacoustics is elucidating the mechanisms that are responsible for combining information across frequencies to form these unified percepts.

Although there are many different approaches one may take to study this process, early experiments measured performance for the detection of a single tone in noise and compared that performance with the detection of more than one tone embedded in noise (Green, 1958; Marill, 1956; Schafer & Gales, 1949). Today, it is generally accepted that the detection of long-duration, independently represented¹ tones follows Green's energy model (Green, 1958). This model posits that the detection of two or more tones is based on a linear combination of the detectability of the individual tones. Detection is predicted to improve with increasing number of tones following $d = \sqrt{\sum_N d_N^2}$, where N is the number of tones. Essentially, such a model is equivalent to statistical summation of independent random variables generated by each tone (e.g., Tanner Jr., 1956).

In a forced-choice experiment, Green's energy detection model predicts a threshold improvement of $10 \log \sqrt{N} = 5 \log N$, or a 1.5 dB improvement for each doubling of N . As mentioned earlier, most empirical studies demonstrate that human listeners tend to achieve such improvements (Buus, Schorer, & Zwicker, 1986; Green, McKey, & Licklider, 1959; Hicks &

Buus, 2000; van den Brink & Houtgast, 1990a, 1990b) for long-duration sounds. For example, van den Brink and Houtgast (1990a, 1990b) measured spectral integration for 100 ms signals and found that as stimulus bandwidth increased, detection thresholds improved at rates of about 1.5 dB per doubling of bandwidth. Using pure tones, Hicks and Buus (2000) measured psychometric functions for the detection of long-duration tones and also found similar rates of integration. In sum, it is observed that by using threshold and accuracy-based measures, spectral integration can follow a linear combination of independent random variables.

One notable characteristic of the aforementioned studies is that tone detection performance was always measured in background noise. That is, the tones were presented to the subjects at relatively high stimulus levels, but the background noise was also at a fairly high level in order to achieve performance below 100%. Yet much of everyday listening is accomplished in fairly low-noise environments, and humans frequently listen in situations where listening is “easy.”

That is, we generally avoid communicating in situations where we might be presented with near-threshold cues. Thus without changing the detection task to one more akin to discrimination (e.g., measuring sensitivity to across-frequency changes), we are unable to measure spectral integration of stimuli presented at suprathreshold levels using accuracy alone. Accuracy-based techniques are also inherently limited when issues pertaining to dynamic mechanisms are raised, because by definition they ignore temporal features of the system and correlate data (e.g., see van Zandt & Townsend, 2013).

However, we are not restricted to evaluating only accuracy in the study of spectral integration, and a more powerful argument can be made if converging evidence is found from multiple techniques. We can apply the *redundant signals approach*, which focuses on multiple signals versus a single signal (e.g., two tones versus one tone) and primarily uses response time (RT) for its dependent variable. The strong tools within that approach can assist an investigator in unveiling the dynamics of the underlying perceptual system and can provide additional tests to evaluate the detection of one versus two signals. We suggest that the two basic measures, accuracy and RT, can together go a long way in answering fundamental questions within the auditory domain. Yet the number of studies applying these RT approaches to the auditory modality is almost nil.

The fundamental goal of this study is to apply the redundant signals techniques to further our understanding of the mechanisms responsible for integrating information across frequency. Green’s model suggests potential architectures for integration of multiple sources of information, and the task sets up the stopping rule constraints (defined in more detail later in the chapter) within which the mechanisms will be tested.

The aforementioned review indicates that spectral integration for long-duration tones follows an optimal and linear integration rule defined by

signal detection theory. An analogous way of thinking about this is an architecture that assumes that the tones are processed by independent channels, and these variables are integrated, or combined, before a decision is made. However, we are unable to evaluate the nature of the architecture using the accuracy methods described previously. Although it is typical and intuitive in the auditory realm to assume that the architecture is parallel for the simultaneous presentation of two tones (see Durlach, Braida, & Ito, 1986, for a commonly assumed parallel model), the same performance scores are often achievable with various different architectures.

Further, there are multiple variants of architecture related to parallel models (as will be discussed in this chapter), and the literature on accuracy is insufficient to distinguish between these models. RT techniques can allow an elucidation of the architectural network for spectral integration. Combined with the large foundation of accuracy data, this complementary approach may greatly refine and improve the models that exist for the detection of complex sounds by the auditory system.

As mentioned earlier, there are limited data available that have applied RT measurements to the detection of one versus many tones. In one monaural study of note, Wagner, Florentine, Buus, and McCormack (2004) demonstrated that broadband noise elicited a faster RT than narrowband noise when measured at the same sound pressure level. Because Wagner et al.’s primary goal was to establish whether RT was determined by loudness rather than sound pressure level, it is ultimately difficult to make a quantitative assessment of the rate of spectral integration. However, as increasing the stimulus bandwidth increases both loudness and number of auditory channels stimulated, their study may be the first to relate RT to spectral integration. The remaining few studies in existence have been accomplished in the binaural domain, where one tone is presented to one ear and a second tone presented to the other. Consequently, we will review the literature on the binaural detection of multiple sounds, but with the knowledge that the complex interactions occurring within the central auditory system may not generalize to detection of tones presented to the *same* ear.

When two tones are presented together (and the two ears are stimulated simultaneously—a binaural presentation), mean RTs are faster than for single ear (or single-tone) detection. Chocholle first demonstrated this result in 1944, and later Simon (1967) demonstrated that the difference in mean RT between binaural and monaural stimulation was very small (about 4 ms for an average 200 ms RT) but statistically significant. More recently, Schlittenlacher, Ellermeier, and Arseneau (2014) also demonstrated a 5–10 ms binaural advantage over monaural stimulation in RT. Thus we see that detection is faster for stimulation of the two ears, but evaluating mean RTs does not allow strong conclusions regarding the mechanisms responsible for the two-ear facilitation.

Schröter, Ulrich, and Miller (2007) undertook a seminal RT-based study within the domain of redundant signals literature in which RT distributions

were evaluated for detection of a 300 ms, 60 dB sound pressure level (SPL) pure tone presented to the left ear, the right ear, or both ears. Schröter et al.'s study was critical in that it evaluated RT *distributions* rather than mean RTs. Importantly, the distributional approach yields a more complete diagnostic test of integration. Whether the two tones presented to the different ears had identical or different frequencies, there was little evidence for a redundant-signals benefit. That is, although RTs were slightly faster for detecting two tones versus one tone, the increase in RT was less than would be expected under probability summation. Their interpretation is that the two ears may not be able to be treated as independent channels. Whether the same result will occur for stimulation by multiple tones to the same ear is yet to be tested.

Further, Schröter et al.'s study did not explicitly evaluate whether inhibitory mechanisms across the ears could have yielded their result or whether attentional resource limitations were at the source. Both explanations are possible sources for the limited dual-ear benefit. The suite of tools from the theory-driven RT methodology Systems Factorial Technology (SFT) originated by Townsend and colleagues (e.g., Schweickert, 1978; Schweickert & Townsend, 1989; Townsend & Nozawa, 1995; Townsend & Wenger, 2004b) allow a thorough assessment of these two explanations.

This methodology permits the simultaneous assessment of a number of critical information processing mechanisms within the same experimental paradigm. These tools will allow an analysis of resource allocation and interaction between the two ears and also provide for psychophysical assessment in high-accuracy situations.

As mentioned earlier, RTs can be measured under conditions at very high stimulus levels associated with high accuracy, tapping into different locations on the psychometric function. RT measures can provide a complement to accuracy-based measures in our attempt at converging on a unified understanding of the mechanisms responsible for perception. Since the broad suite of tools available in SFT has not heretofore been implemented in auditory perception, except for a single study by Lentz, He, and Townsend (2014), which focused on binaural masking phenomena, the following section provides a brief overview.

7.1.1 Architecture

One of the first issues to address is the form, or the architecture, used by a system. We define *serial processing* as processing things one at a time or sequentially, with no overlap among the successive processing times. *Parallel processing* means processing all things simultaneously, although it is allowed that each process may finish at different times (e.g., Townsend & Ashby, 1983; Townsend, Yang, & Burns, 2011). *Parallel coactive processing* refers to a system in which the channels are summed together before a

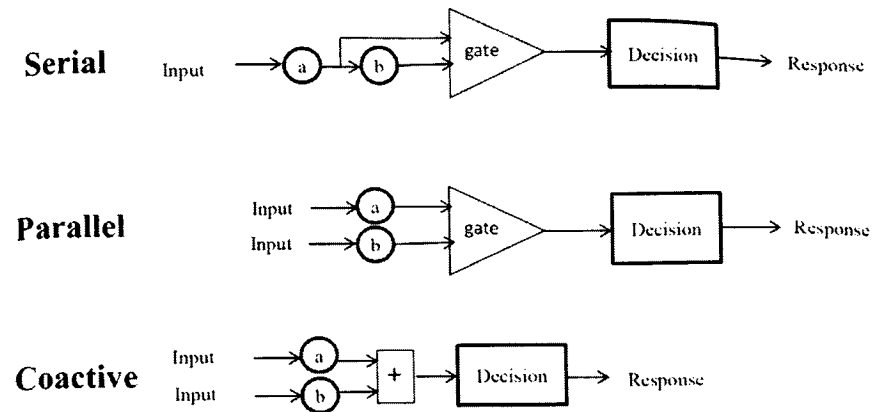


Figure 7.1 Illustration of three models: serial, parallel, and coactive.

decision is made (Haupt & Townsend, 2011; Townsend and Nozawa, 1995). These three models are illustrated graphically in Figure 7.1.

For decades, the almost universal practice in testing serial versus parallel processing has been to vary the processing load (i.e., number of inputs, n), and then to plot the slopes of the mean response times as a function of n . If the slope of such a graph differs significantly from zero, then processing is declared to be serial. If it does not differ significantly from zero, parallel processing is inferred. This reasoning is fallacious on several grounds (see Townsend, 1990; Townsend et al., 2011), and although thorough discussion is beyond the scope of this chapter, more sophisticated techniques are needed to assess architecture than simply varying the processing load.

Sternberg's celebrated additive factors method (Sternberg, 1969) offered a technique that avoided the aforementioned conundrum and could affirm or deny serial processing. The new method, which Sternberg viewed as a natural extension of Donders's (1869) method of subtraction, was based on the notion of "*selective influence*" of mean processing times, which stipulated that each experimental factor affects one and only one psychological subprocess at the level of means. The challenge there was that the method did not directly test other important architectures such as parallel systems. Also, there was a lack of mathematical proof for the association of "factors that are additive" even with serial processing if the successive times were not stochastically independent and, again, there was no clear way to include other architectures.

As we will discuss in the subsequent section in more detail, Townsend and colleagues proved that if selective influence acted at the level of distributional ordering, many different architectures, including parallel and

serial ones, could be discriminated at the level of mean RTs (Townsend, 1984, 1990; Townsend & Ashby, 1983; Townsend & Schweickert, 1989). Selective influence at the level of the RT distributions provides even stronger inferences (Townsend & Nozawa, 1995; Townsend & Wenger, 2004a).

However, we also must discuss *independence* versus *dependence* of channels, stages, or subsystems (these terms can be used interchangeably, although the term stages is sometimes restricted to serial systems and channels to parallel systems). In serial processing, if the successive items are dependent, then what happens on *a*, say, can affect the processing time for *b*. Although it is still true that the overall mean exhaustive time will be the sum of the two means, the second, say *b*, will depend on *a*'s processing time. In a parallel system, speeding up *a* could either speed up or slow down *b* because they are being processed simultaneously; inhibition or facilitation (or both) can take place during a single trial, while processing is ongoing. Townsend and Wenger (2004b) discuss this topic in detail, and interactions between channels may occur readily within the auditory system. Unfortunately, strong interactions between channels could manifest themselves in a loss of measureable selective influence.

7.1.1.1 Stopping or Decision Rule: When Does Processing Cease?

No predictions can be made about processing times until the model designer has a rule for when processing stops. In some high-accuracy situations, such as search tasks, it is usually possible to define a set of events, any one of which will allow the processor to stop without error. In search of a set of targets, then, the detection of any one of them may in some cases serve as a signal to cease processing.

A special case ensues when exactly one sought for target is present. In any task where a subset of the display or memory items is sufficient to stop without error, and the system processor is capable of stopping (not all may be), as noted earlier, the processor is said to be capable of *self-termination*. Because many earlier investigations (e.g., Sternberg, 1966) studied exhaustive versus single-target search, self-termination was often employed to refer to the latter, although it can also have generic meaning and convey, say, *first-termination* when the completion of any of the present items suffices to stop processing. The latter case is often called an OR design because completion of any of a set of presented items is sufficient to stop processing and ensure a correct response (e.g., Egeth, 1966; Townsend & Nozawa, 1995).

The OR design can be visualized in Figure 7.1 in which an OR operation would be placed at the gate. In the serial example, input *a* could stop processing prior to input *b* being fully processed, as the decision could be made

with either input passing through the OR gate. In the parallel example, either *a* or *b* could stop processing—both inputs need not be received by the decision element for a decision to be made. Notably, one could call this a *race model*, as detection is determined by the first input to reach the decision element.

If all items or channels must be processed to ensure a correct response, then exhaustive processing is entailed. For instance, on no-target (i.e., nothing present but distractors or noise) trials, every item must be examined to guarantee no targets are present. In an experiment where, say, all *n* items in the search set must be a certain kind of target, called an AND design, exhaustive processing is forced on the observer (e.g., Sternberg, 1966; Townsend & Nozawa, 1995). The AND design can also be visualized in Figure 7.1 in which an AND operation would be placed at the gate. In the serial example, input *a* would need be processed first and input *b* would be processed second. In the parallel example, both inputs must reach the decision element. In both cases, a decision cannot be made until processing has been completed for both *a* and *b*.

To some extent, the stopping rule is determined by the experimental design—participants can be asked to respond to either input *a* or input *b* (an OR task) or to both input *a* and input *b* (an AND task). Nevertheless, as intimated earlier, some systems may by their very design have to process everything in the search set, so the question of the stopping rule is of interest even when, in principle, self termination is a possibility.

Importantly, the techniques developed by Townsend and colleagues (e.g., Townsend, 1984, 1990; Townsend & Ashby, 1983; Townsend & Schweickert, 1989) allow a rigorous analysis of system architecture. Consider an experimental design (commonly referred to as a double-factorial design) in which the two experimental variables (*a* and *b*) are presented at two different levels (1 and 2). The levels should be selected as to manipulate RTs (e.g., level 2 might be associated with a faster RT than level 1). To evaluate architecture, we must present the two variables together at each combination of levels (a^1b^1 , a^1b^2 , a^2b^1 , a^2b^2). Let us define \overline{RT}_{11} as the mean RT for when stimulus *a* is presented at level 1 and stimulus *b* is present at level 1. We then define the *mean interaction contrast* as follows:

$$MIC = \overline{RT}_{11} + \overline{RT}_{22} - \overline{RT}_{12} - \overline{RT}_{21}. \quad (7.1)$$

The value of the MIC (positive, zero, or negative) qualitatively provides for distinguishing between the various serial and parallel models. Subsequent work, and that which we attempted to implement here, extended such theorems to the more powerful level of entire RT distributions (Townsend & Nozawa, 1995; Townsend & Wenger, 2004b). Rather than simply using mean RTs, this work uses the full distributions characterized as

Table 7.1 Summary of mean and survivor interaction contrast propositions.

	MIC	SIC
Self-terminating parallel	Positive	Positive
Exhaustive parallel	Negative	Negative
Self-terminating serial	Zero	Zero
Exhaustive serial	Zero	Negative going to positive
Channel summation	Positive	Negative going to positive

survivor functions, which are determined from the cumulative distribution functions ($S(t) = 1 - F(t)$). The *survivor interaction contrast* is defined as follows and can be evaluated for all RTs:

$$SIC(t) = S_{11}(t) + S_{22}(t) - S_{12}(t) - S_{21}(t). \quad (7.2)$$

We can use the MIC (Equation 7.1) and the more rigorous SIC functions (Equation 7.2) to evaluate the architecture as long as we have evidence for selective influence. Table 7.1 shows the general patterns of MIC and SIC functions for the various architectures, whereas Figure 7.2 plots the patterns of SIC functions expected.

7.1.2 Capacity and Workload Capacity: Various Speeds on a Speed Continuum

Capacity generally refers to the relationships between the speeds of processing in RT tasks. Workload capacity will refer to the effects on efficiency as the workload is increased. For greater mathematical detail and in-depth discussion, see Townsend and Ashby (1978), Townsend and Nozawa (1995), and Townsend and Wenger (2004a). Wenger and Townsend (2000a) offer an explicit tutorial and instructions on how to carry out a capacity analysis.

Informally, the notion of *unlimited capacity* refers to the situation when the finishing time of a subsystem (item, channel, etc.) is identical to that of a standard parallel system; that is, the finishing times of the distinct subsystems are parallel, and the average finishing times of each do not depend on how many others are engaged. *Limited capacity* refers to the situation when item or channel finishing times are less than what would be expected in a standard parallel system. *Super capacity* indicates that individual channels are processing at a rate even faster than standard parallel processing. Figure 7.3 illustrates the general intuitions accorded these concepts, again in an informal manner. The size of the cylinders provides a description of the amount of resources available.

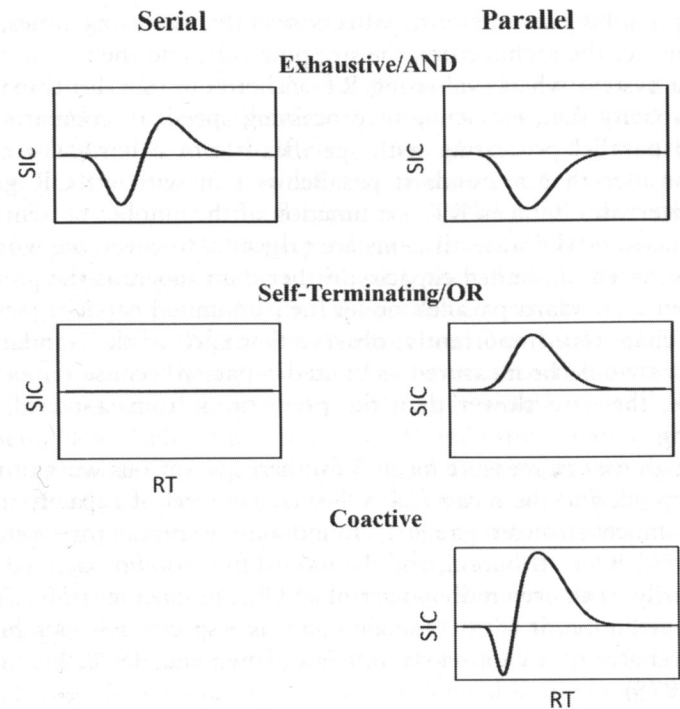


Figure 7.2 Predicted shapes of SIC functions for the various forms of architecture and stopping rules. Serial systems are shown in the left panels and parallel systems in the right. The upper panels depict AND predictions and the middle panels depict OR predictions, with the bottom right panel showing coactive predictions.

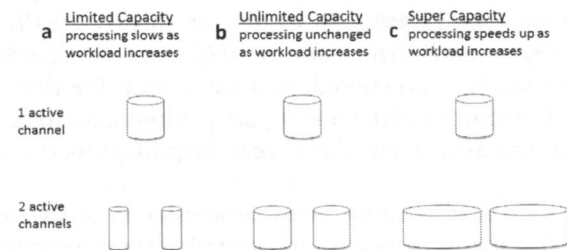


Figure 7.3 Graphical intuition of a system's behavior under different capacity bounds: unlimited capacity, limited capacity, and super capacity.

The stopping rule obviously affects overall processing times, so one must consider the architecture, the stopping rule, and the resource allocation of a system when evaluating RTs. Therefore, our benchmark is to assess capacity (i.e., efficiency of processing speed) in comparison with standard parallel processing with specification of a particular stopping rule. Consider that a standard parallel system with an OR gate will predict decreases in mean RT as a function of the number of items undergoing processing (because all items are targets). However, we would consider this system unlimited capacity, rather than super, as the predictions arise from a standard parallel model (i.e., unlimited capacity with independent channels). Importantly, observe that each of the standard serial predictions would be measured as limited capacity because for each stopping rule, they are slower than the predictions from standard parallel processing.

Although we can measure mean RTs, there are various ways of measuring this speed, and the mean is a rather coarse level of capacity measurement. A much stronger gauge is found in the cumulative distribution function $F(t)$ RT distribution, and the hazard function $h(t)$ (to be discussed momentarily) is an even more powerful and fine-grained measure. Ordering on the distribution or hazard-function level is a special case of a hierarchy on the strengths of a vital set of statistics (Townsend, 1990; Townsend & Ashby, 1978).

The ordering establishes a hierarchy of strength because, say, if $F_a(t) > F_b(t)$, then the mean of a is less than the mean of b . However, the reverse implication does not hold (ordered means do not imply ordering of the cumulative distribution functions). Thus ordering at the distributional level is much more powerful than ordering of the means. Obviously, if the cumulative distribution functions are ordered, then so are the survivor functions ($S(t) = 1 - F(t)$). That is, $F_a(t) > F_b(t)$ implies $S_a(t) < S_b(t)$. Similarly if $h_a(t) > h_b(t)$, then $F_a(t) > F_b(t)$, but not vice versa, and so on.

There is a useful measure that is at the same strength level as $F(t)$ or $S(t)$, as it is a monotonic transformation. This measure is defined as $-\ln S(t)$. Wenger and Townsend (2000a) illustrate that this is actually the integral of the hazard function $h(t')$ from zero to t ($H(t)$ e.g., Wenger & Townsend, 2000b; see also Neufeld, Townsend, & Jetté, 2007). We thus write the integrated hazard function as $H(t) = -\log[S(t)]$. Although $H(t)$ is of the same level of strength as $S(t)$, it has some very helpful properties not directly shared by $S(t)$.

It has been demonstrated that when processing is of this form, that is, standard parallel, the sum of the integrated hazard functions for each item presented alone is precisely the value, for all times t , of the integrated hazard function when both items are presented together (Townsend & Nozawa, 1995). That is, $H_a(t) + H_b(t) = H_{ab}(t)$. This intriguing fact suggests the formulation of a new capacity measure, which Townsend and Nozawa (1995) called the *workload capacity coefficient* $C(t) = H_{ab}(t)/[H_a(t) + H_b(t)]$;

that is, the ratio of the double item trials over the sum of the single target trials.

If this ratio is identical to one for all t , then the processing is considered *unlimited*, as it is identical to that of an unlimited capacity-independent parallel model. If $C(t)$ is less than one for some value of t , then we call processing *limited*. For instance, either serial processing of the ordinary kind or a fixed-capacity parallel model that spreads the capacity equally across a and b predicts $C(t) = 1/2$ for all times $t > 0$. If $C(t) > 1$ at any time (or range of times) t , then we call the system *super capacity* for those times. A tutorial on capacity and how to assess it in experimental data is offered in Wenger and Townsend (2000a). In a recent extension of these notions, it has been shown that if configural parallel processing is interpreted as positive interactive parallel channels (thus being dependent or positively correlated rather than independent), then configural processing can produce striking super capacity (Eidels, Houpt, Altieri, Pei, & Townsend, 2011; Townsend & Wenger, 2004a).

Subsequently, a general theory of capacity was formulated that permitted the measurement of processing efficiency for all times during a trial (Townsend & Nozawa, 1995; Townsend & Wenger, 2004b). Employing standard parallel processing as a cornerstone, the theory defined unlimited capacity as efficiency identical to that of standard parallel processing in which case the measure is $C(t) = 1$. It defined limited capacity as efficiency slower than standard parallel processing. For instance, standard serial processing produces a measure of capacity of $C(t) = 1/2$. Interestingly, that is the same prediction as for a limited capacity parallel system, which divides all the capacity that was present for $N = 1$ across all the channels for $N > 1$, a so-called fixed capacity parallel system. And finally, the theory defined super capacity as processing with greater efficiency than standard parallel models could produce; that is, $C(t) > 1$.

In sum, our measuring instrument is that of the set of predictions by unlimited-capacity independent parallel processing (UCIP). As mentioned earlier, *unlimited capacity* means here that each parallel channel processes its input (item, etc.) just as fast when there are other surrounding channels working (i.e., with greater n) as when it is the only channel being forced to process information. The purpose of this paper is to apply these techniques, with a focus on evaluating the architecture for the detection of one versus two tones and the allocation of resources.

To most closely follow the experiments set forth by Green (1958) and subsequent replications, we adopted an OR task here, which required a participant to push a button when a sound was heard. Thus a participant would select "yes" if sound a , sound b , or sound $a + b$ was presented. By defining the experiment in this way, we can look for evidence of serial, parallel, or coactive processing using the architecture tools. We also assess resource limitations by measuring capacity.

7.2 Methods

7.2.1 Stimuli

As this study is the first robust implementation of SFT techniques to the auditory system, it was critical that we design stimuli with a number of constraints in mind. Most importantly, we wanted to ensure that the two tones were independently represented in the auditory system. Consequently, the frequencies selected are chosen to be well-spaced in frequency (in this case, more than two octaves apart: 500 and 3,020 Hz).

Stimuli were 350 ms duration, 500 and 3,020 HZ pure tones having 25 ms cosine-squared onset and offset ramps. These target tones were presented at two stimulus levels: 38 and 80 dB SPL. Although it may seem that these two stimulus levels are widely different, these levels were chosen such that distributional ordering would be observed for single-target trials; that is, the higher-level stimulus would be detected more quickly than the lower-level stimulus. We expected around a 50 ms difference in RT for the two different levels used here (Kohfeld, 1971).

7.2.1.1 Procedures

Four observers participated in experimental sessions lasting one to two hours. A single session consisted of 6–12 blocks of 128 trials. Each trial began with a visual warning of “listen” appearing on a computer monitor for 400 ms. A random-duration, uniformly distributed silent period ranging between 0 and 400 ms followed removal of the warning, when the trial began.

Following a standard double-factorial design, on each trial there were two possible events: A stimulus or a silent interval. Within a single block of 128 trials, 75% of the trials were stimulus trials (“Yes” trials): 25% of total trials contained only the 500 Hz tone, 25% contained only the 3,020 Hz tone, and 25% contained both 500 and 3,020 Hz tones. When single tones were presented, the High and Low stimulus levels were presented in equal proportions. When tones are presented together, four possible, equally probable subevents arise: High-High (HH), both tones at the High level; Low-Low (LL), both tones at the Low levels; High-Low (HL), with the 500 Hz tone at the High level and the 3,020 Hz tone at the Low Level; and Low-High (LH), with the 500 Hz tone at the Low level and the 3,020 Hz tone at the High level. See Table 7.2 for a depiction of the various events. On the remaining 25% of trials, no stimulus (NS) was presented (“No” trials).

Two conditions had both tones presented either to the left ear or to the right ear.² Trials were presented using a randomized block design in which the condition was selected randomly, and three blocks were run in that condition. Then, a new condition was selected and three blocks were again tested. Once the different condition types were tested, the order of the

Table 7.2 A list of the various possible events in the double-factorial design. Each row represents an occurrence with frequency of 1/16th. Seventy-five percent of the trials are “yes” trials in which a stimulus was presented. Twenty-five percent of the trials are “no” trials in which no stimulus was presented (indicated by NS).

	500 Hz	3020 Hz	
Dual targets “Yes”	High	High	HH
	High	Low	HL
	Low	High	LH
	Low	Low	LL
Single Targets “Yes”	High	NS	H500
	High	NS	H500
	Low	NS	L500
	Low	NS	L500
	NS	High	H3020
	NS	High	H3020
	NS	Low	L3020
	NS	Low	L3020
Target absent “No”	NS	NS	
	NS	NS	
	NS	NS	
	NS	NS	

conditions was again randomly selected, and all blocks were retested three times. A total of 12 blocks were collected for each condition, yielding a total of 96 trials for each dual target (HH, LL, HL, and LH) and 192 trials for the single targets (H500, H3020, L500, and L3020).

Stimuli were presented to the observers at a 24,414 kHz sampling rate using a 24-bit Tucker Davis Technologies (TDT) RP2.1 real-time processor. We digitally summed the two tones at appropriate relative levels prior to playing them through a single channel of the RP2.1. We then passed the signals through a calibrated PA5 programmable attenuator and an HB6 headphone buffer connected to a Sennheiser HD280 Pro headphone set, which was used to present signals to the listeners. We measured the RTs using a button box interfaced to the computer through the TDT hardware.

Four listeners, ranging in age from 19 to 44 participated in the experiment. All subjects had hearing thresholds of 15 dB HL or better in both ears at all audiometric frequencies. Obs. 4 is the first author.

Observers were instructed to respond as quickly to the signal tone as possible while attempting to provide correct responses. Using an “OR” design, observers were required to respond with the “yes” button if a tone was present. Otherwise, they were instructed to respond with the “no” button. The RT was measured from the onset of the tone stimulus. Percent correct was monitored in order to ensure that subjects achieved

high levels of performance.³ Note that we removed RTs below 80 ms or greater than 800 ms from the data set.

7.3 Results

7.3.1 Single Targets

We first report the RT distributions to familiarize the reader with the data format and to present the robust RT distributional data. Figure 7.4 plots survivor functions for the high- and low-stimulus levels for the two frequencies presented to the left and right ears for a single subject. Recall that the survivor function $S(t)$ is simply $1 - F(t)$, where $F(t)$ represents the cumulative distribution function of RTs. We present data from a representative single subject (Obs. 2) because of the overwhelming similarity in the pattern of results across the subjects.

As one can see from Figure 7.4, there is little overlap in the distributions obtained for the High- and Low-stimulus intensities. Notably, all subjects demonstrated significantly faster RTs for the High intensity versus the Low intensity at both frequencies and for both ears. For all subjects at both frequencies and for both ears, Kolmogorov-Smirnov (KS) tests of survivor function orderings revealed significance well below the $p < 0.0001$ level between the high- and low-RT distributions. Consequently, the data present a compelling case that distributional ordering is present for single targets. Although some subjects showed one ear being faster than the other or one frequency being faster than the other at the distributional level, there were no systematic patterns across the four subjects in this regard.

Mean RTs are also listed in Table 7.3 to demonstrate the similarity in patterns of RTs across the subjects. We show geometric means due to the skewness of the RT distributions. On average, subjects are 43 ms faster at the High level than at the Low level.

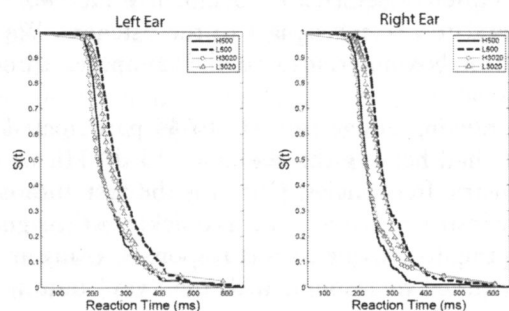


Figure 7.4 Survivor functions for the different types of single-target trials. Data from Obs. 2 reported.

Table 7.3 Mean (geometric) RTs in milliseconds for single targets.

	Left				Right			
	500 Hz		3020 Hz		500 Hz		3020 Hz	
	High	Low	High	Low	High	Low	High	Low
Obs. 1	182.7	245.9	180.9	221.0	183.0	254.8	181.1	250.4
Obs. 2	205.4	237.8	194.7	222.2	203.8	239.8	197.6	227.2
Obs. 3	180.1	216.1	174.1	203.7	184.1	218.0	180.4	216.6
Obs. 4	244.0	299.7	238.4	278.2	238.2	294.4	237.8	312.2
Mean	201.5	248.1	195.5	229.7	201.1	250.3	197.9	249.0

Table 7.4 Mean (geometric) RTs in milliseconds for the dual-target trials.

	Left Ear				Right Ear			
	HH	HL	LH	LL	HH	HL	LH	LL
Obs. 1	211.8	228.7	215.7	258.3	209.3	216.6	224.9	286.0
Obs. 2	227.4	237.9	239.1	266.3	217.3	239.8	233.7	254.2
Obs. 3	197.7	211.2	203.0	223.9	199.4	209.4	209.2	233.0
Obs. 4	266.1	264.5	274.8	319.0	261.0	267.9	273.2	325.3
Mean	224.4	234.8	231.6	264.8	220.6	232.3	234.1	272.4

7.3.2 Dual Targets

Figure 7.5 plots the estimated survivor functions for the dual target data for the left ear (left panels) and for the right ear (right panels) for all subjects, and Table 7.4 reports the mean RT values for each of the trial types. When evaluating the mean RTs for all subjects, there is a strong ordering with $HH < HL \approx LH < LL$. Figure 7.5 also provides some support for distributional ordering at the individual level, but only Obs. 2 and 3 appear to have observable patterns of distributional ordering. This pattern is evident with survivor functions for HH being toward the left of the panels and LL being toward the right, with HL and LH (indicated by symbols) in between. However, Obs. 1's right ear and Obs. 4's left do not appear to have a strong ordering, as the HH and HL functions cross in the middle of the distributions. To establish whether there were statistically significant distributional orderings, KS tests of survivor function orderings were conducted and are reported in Table 7.5.

Table 7.5 p values for one-tailed KS tests for survivor function orderings obtained from the dual targets. Bolded values indicate statistical significance at the $p < 0.05$ level. HH was always faster than LL below the $p < 0.0001$ level. The ** indicates subjects for which the dominance tests pass for all comparisons.

	Left Ear				Right Ear			
	Obs. 1	Obs. 2	Obs. 3**	Obs. 4	Obs. 1	Obs. 2	Obs. 3**	Obs. 4
HH < LH	0.47	0.17	0.048	0.37	0.27	0.15	0.026	0.13
HH < HL	0.018	0.025	0.001	0.28	0.087	0.0014	0.019	0.11
LL > LH	<	<	0.0001	<	<	0.0040	<	<
LL > HL	0.0001	0.0001		0.0001	0.0001		0.0001	0.0001
LL > LH	0.0023	<	0.01	<	<	0.29	<	<
LL > HL		0.0001		0.0001	0.0001		0.0001	0.0001

Table 7.5 demonstrates statistically significant survivor function orderings for all stimulus levels for Obs. 3 only. Data from the other subjects do not meet the criterion for statistically significant ordering, with a failure of the test in at least one, but sometimes two, comparisons.

There are many possible reasons for the lack of significant selective influence observed for three of four observers here. One factor may be insufficient statistical power to establish differences between the different RTs. More trials may be needed to identify statistically significant distributional orderings. As the mean RTs demonstrate, the difference in mean RT between the HH and LH and HL targets can be relatively small—around 8–10 ms. Notably, these differences are approximately a factor of 1.5 or 2 smaller than observed in Townsend and Nozawa (1995). For some of the cases (e.g., Obs. 2 both ears, Obs. 1 right ear, and Obs. 4 left ear), the functions look ordered, but they are very close to each other. Thus we can entertain the possibility that more trials might have allowed a statistically significant result. However, we should also consider that factors specific to the auditory system and the task could also contribute to this finding. In particular, Obs. 4's left ear and Obs. 1's right ear show an ordering of the means, but Figure 7.5 demonstrates that the survivor functions for HL crosses those of HH. This is a more robust failure of selective influence, as it would be very unlikely to occur because of the insufficient number of trials. Such a failure is more likely to arise from interactions occurring between the two channels or perceptual factors. One final possibility is that of ceiling effects: Listeners are very fast in these tasks (which is fairly common for RTs in the auditory domain). However, it is possible

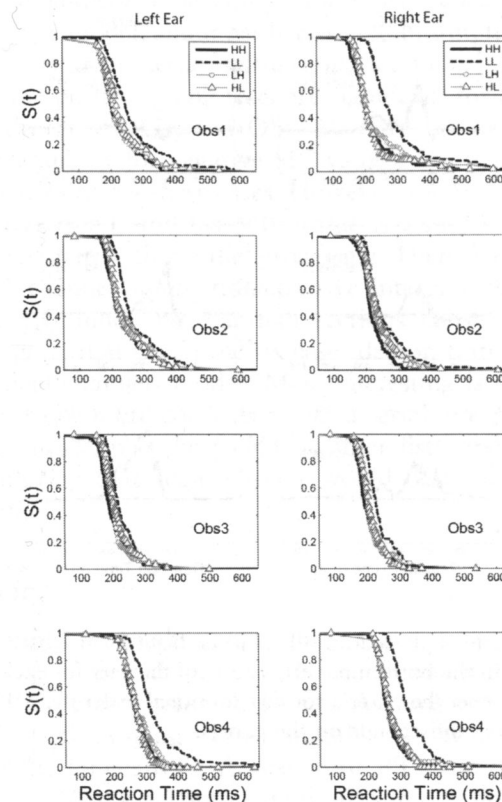


Figure 7.5 Survivor functions for the four subjects for the dual-target trials.

that HH simply cannot be any faster than either HL and LH, as those RTs may already be at ceiling.

7.3.3 Architecture

Although not all subjects meet a highly conservative definition of selective influence (via a significant KS test result), in most cases, we observe ordering of the survivor functions. Consequently, we argue that distributional ordering is present in six of eight cases, but that it is not detectable by our statistical tests for all subjects. We therefore analyze the architecture for the subjects in which we believe that distributional ordering is present.

Figure 7.6 plots the SIC functions obtained using Equation 7.2 for the left and right ears with MIC values in milliseconds (obtained from Equation

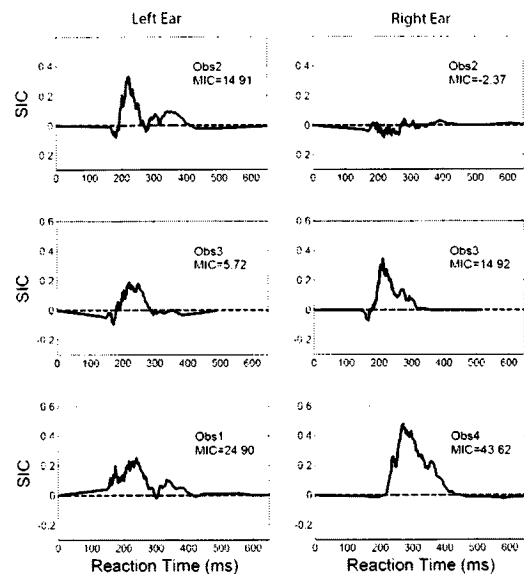


Figure 7.6 SIC functions plotted for all subjects. Data from Observers 1 and 4 are plotted in the bottom panels, as one of the ears for each of the subjects did not meet the criteria for distributional ordering. MIC values are reported in milliseconds on the panels.

7.1) in the upper right-hand corner of each panel. Generally, all functions show a positive trend except for Obs. 2's right ear, which hovers around zero. The MIC values for all subjects are positive, except for this one case, where the $MIC \approx -2$. There are also some cases where there is a small negative dip at the short RTs. Recall from Figure 7.2 that there are three expected shapes for SIC functions for self-terminating stopping rules (OR): zero (serial), positive (parallel), and negative going positive (coactive). Except for Obs. 2's right ear, observation suggests that the architecture is either parallel or coactive.

To establish whether there is support for the parallel architecture, we conducted statistical tests using Houpt and Townsend's (2010) formulation. There are two sides to this test: one tests whether the largest positive SIC value is different from zero and the other tests whether the largest negative SIC value is different from zero. A significant positive test indicates that we can reject serial, self-terminating models and parallel, exhaustive models. A significant negative test indicates that we can reject serial, self-terminating models and parallel self-terminating models. If both are

significant, we would conclude support for either a coactive architecture or a serial, exhaustive model, and if neither are significant, the interpretation is that we cannot reject the serial, self-terminating model.

In general, the positive going tests are significant for all subjects except Obs. 3's left ear ($p = 0.17$), and Obs. 2's right ear ($p = 0.92$). None of the tests are significant for negative SIC values. Consequently, we cannot reject serial processing for these cases. However, the SIC for Obs. 3 demonstrates a positive trend, and his MIC value is also positive (as shown in Figure 7.6), consistent with parallel processing. There was no significant evidence for the channel summation/coactive interpretation as no subjects demonstrated SIC functions that had sections consistently below zero. Overall, we suggest that five of the six cases demonstrate support for a parallel, independent-channels model. Most interesting is Obs. 2's right ear, which appears to be more consistent with a serial interpretation. Whether this result will hold up or be found in other listeners is yet to be seen, and it is perplexing that such a result would occur for one of her ears and not the other.

7.3.4 Capacity

Capacity functions for the left and right ears are shown in Figure 7.7. All capacity values are plotted on the same panel, and different symbols are used for the different salience levels. All subjects show a similar pattern, with $C(t)$ ranging from roughly 2.0 at the fastest RTs to around $\frac{1}{2}$ for the longer RTs. Across most RTs, the results consistently demonstrate $1 \geq C(t) \geq \frac{1}{2}$. Both ears of all subjects demonstrate statistically significant limited capacity $C(t) < 1$ by the Houpt and Townsend (2012) statistics.

All subjects also have small RT ranges (for the short RTs only) in which capacity values tend to exceed 1.0. When examining the point-wise confidence intervals around the capacity values based on the individual subjects, we see no evidence that $C(t)$ is significantly greater than 1.0 for any RT. Thus we do not conclude super capacity for any of the subjects tested here. Generally speaking, though, the trend is for capacity values to range from unlimited capacity to limited/fixed capacity, with a dominance of limited capacity.

Grice and colleagues proposed a lower bound on performance parallel systems (e.g., Grice, Canham, & Gwynne, 1984) that applies to limited capacity. In this case, performance on double-target trials is slower than on those single-target trials that contain the faster of the two targets. When performance on the two channels is equal, the Grice bound indicates efficiency at the level of *fixed capacity* in a parallel system (i.e., $C(t) = \frac{1}{2}$). A fixed capacity system can be viewed as sharing a fixed amount of capacity between the two channels. Alternatively, a serial system can make exactly this prediction as well (Townsend & Wenger, 2004a). If the Grice

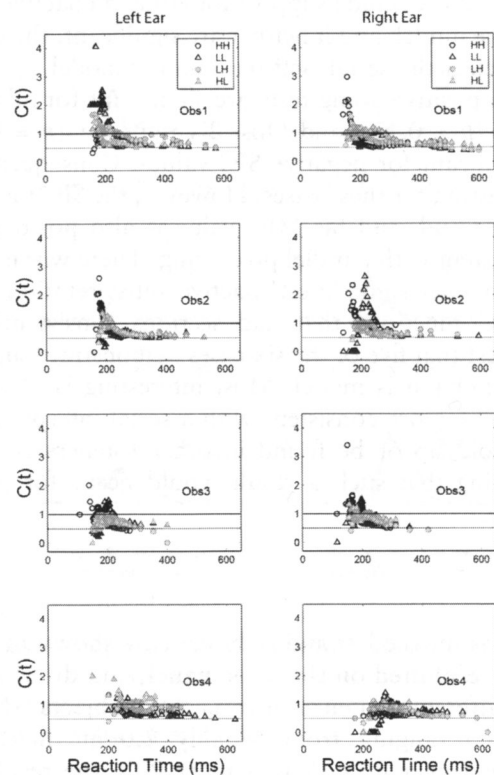


Figure 7.7 Capacity values for the left and right ear at the four different salience levels.

inequality is violated (that is, capacity values fall below the bound), the system is limited capacity in a very strong sense (Townsend & Wenger, 2004a). If the RTs for the two frequencies were the same, the Grice bound = $\frac{1}{2}$ in the capacity space. Although it is possible to calculate the different Grice bounds for each of the four functions within the panel (see Townsend & Eidels, 2011), we use $C(t) = \frac{1}{2}$ as a general estimate of the Grice bound so as to not clutter the figures. The empirically calculated Grice bound is very close to one-half for all RTs, although sometimes slightly higher.

For most RTs, capacity tends to be at or slightly better than one-half. We see that the vast majority of the capacity values are above the Grice bound, and very few capacity values that violate (fall below) the Grice bound. We interpret this result as a redundancy gain—listeners benefit from (that is, RT is speeded up by) the presence of two tones versus a single tone.

7.4 General Discussion

Heretofore, as far as we are aware, only para-threshold, accuracy experiments have investigated the ability to detect multiple stimuli within a single ear. This study presents the first study in which RT tools have been used to evaluate the architecture of monaural auditory perception and the application of capacity measures. We will first discuss general results for the single-target and dual-target trials as they relate to the published literature. Then, we will evaluate the results put forth by this experiment in light of the energy model posited by Green (1958) and use these results to inform these accuracy-based models.

The single target data represent robust distributional data not previously reported in the psychoacoustic literature, but we can also compare the mean RTs obtained here with previously published studies. The experimental factor, stimulus intensity, was effectual in strongly ordering the single-target survivor functions. For the single-target RTs, we obtained very similar mean RT differences to those reported in the literature: approximately a 43 ms difference between the high- and low-stimulus intensity signals (Kohfeld, 1971).

We also found very similar RTs for the 500 and 3,020 Hz stimuli. Such a result is consistent with previous studies linking the loudness of stimuli to their RT (Marshall & Brandt, 1980; Wagner et al., 2004). When evaluating equal-loudness contours (Fletcher & Munson, 1933), we see that 500 and 3,020-Hz have loudness levels similar to each other at the high and low levels used here. Finally, we did not observe a difference in mean RT between the left and right ears, consistent with that observed by Schlittenlacher et al. (2014).

We also observed that stimulus intensity ordered the dual-target survivor functions, but only for one of the four observers. Failure of ordering tended to be most common between HH and LH, although we did also see failures between HH and HL. As can be seen from Table 7.4, there is a clear ordering of the means $HH < HL \approx LH < LL$, but this ordering did not hold at the level of the survivor functions for three of four observers. We attempted to achieve selective influence of the dual targets by using tones that are well-separated in frequency (e.g., they would be independently represented by the auditory system), and it seems likely that further frequency separation would not increase the likelihood of obtaining selective influence, as the tones are already separated by a factor of over six. It may be possible that more stimulus trials would garner greater statistical power, but this may not be the problem. Note that tones presented at high supra-threshold stimulus levels may be interacting within the auditory system at a peripheral level (see Egan & Hake, 1950). We could test this possibility by reducing the overall levels of the stimuli and observing whether distributional ordering is present for lower-level stimuli, which are less likely to interact within the ear. We may also

have encountered ceiling effects: If HL and LH were already associated with the fastest possible RT achievable by the subject, HH could not possibly have been faster.

Although the random delay we imposed upon the onset of the stimuli was intended to avoid ceiling effects and anticipation responses, it is possible that subjects were not sufficiently slowed by that delay.

For the presentation of multiple simultaneous tones to the same ear, a common assumption in the psychophysical literature is a parallel architecture (e.g., Durlach et al., 1986; Florentine & Buus, 1981). It does make intuitive sense that multiple tones presented simultaneously would require a parallel architecture for their perception. However, this assumption appears to be made without strong experimental support, and accuracy-based measures are unable to assess differences between serial and parallel processing (e.g., Eidels, Townsend, Hughes, & Perry, 2015). For accuracy-based evaluations, there are many circumstances where parallel and serial models could give identical predictions. The current study may very well be the first one to demonstrate a parallel self-terminating architecture for the detection of multiple tones. For the majority of subjects, we have support for a parallel self-terminating architecture, as the SIC functions were positive and not negative. Yet it is noteworthy that one subject demonstrated strong support for serial processing. Should this finding hold up for additional subjects, it would put into question some of the basic tenets of the models used to describe/explain psychoacoustic data.

We next consider models that have been applied to low-to-moderate-accuracy monaural tone detection. The most prominent is David Green's (1958) energy model (here referred to as the "integration model"), which we discuss presently, and a model that he employed to test against the energy model, a model he termed the "independent thresholds model." We immediately run into an important factor that distinguishes most accuracy-based from most RT-based models in psychophysics: The former are almost entirely mute with regard to time; the latter, of course, are not. Thus the classical models for detection and identification (e.g., Green & Swets, 1966; Luce, 1963) assume a random sample is taken of the perceptual representation and a consequent decision is made from that sample.

The independent thresholds model assumes that a sample is garnered from each signal and if either exceeds a threshold associated with its own channel, then a positive detection occurs. If neither signal exceeds its threshold, guessing occurs. Hence this model proposes that detection in the presence of more than one signal follows a probability summation principle.

Green's version is a "high-threshold" model in that it is posited that no false alarm ever occurs on noise-alone trials. Also, because the independent-thresholds model assumes no degradation of performance when two, rather

than one, channels are active, we have the accuracy correspondent to unlimited capacity. The closest analogue to this model in the RT literature seems to be one where decisions are made on the separate channels when the activation reaches a decision criterion in either one, thus delivering unlimited-capacity, probability summation in terms of time rather than accuracy.

On the other hand, Green's energy model posits that the detection of two or more tones is based on a linear combination of the detectability of the individual tones. He picks the weights as equivalent to the means of each variable, a move that produces a matched filter representation, which is an optimal linear filter under prescribed conditions (Green & Swets, 1966; Townsend & Landon, 1983). Observe that this model also assumes no degradation in the individual channels with an increasing number of signals, but the decision rule is, of course, quite different from the independent-thresholds model.

Detection is predicted to improve with the increasing number of tones following $d' = \sqrt{\sum_N d_N'^2}$, where N is summed over the number of tones presented. A coactive parallel model affords an RT analogue to this linear combination, as the combination of channels occurs prior to the decision variable. Consequently, Green's model, in the accuracy domain, is analogous to a coactive system, thereby producing super capacity in the overall two-signal performance.

Our results support parallel processing but with moderately limited capacity. Although our data only indicate relatively weak selective influence, the conclusions of parallelism are easy to believe. The architecture tests do not evidence the early negative blips characteristic of coactive models (Townsend & Nozawa, 1995), although the relatively low power might make those hard to detect in our data. The capacity analyses are rock solid in favor of moderate-to-low capacity. The fact that coactive models typically predict extreme super capacity (Townsend & Wenger, 2004b) also makes coactivation an unlikely candidate in this study.

At this point, we cannot pinpoint the source of the limited capacity. Potentially, it could simply be the result of limitations in attentional resources. Alternatively, and we suspect more likely, it could be due to something like mutual inhibition across the parallel channels. The fact that even binocular dot experiments reveal limited capacity (e.g., Townsend & Nozawa, 1995) suggests that certain kinds of interaction or restricted sources of processing facility can intrude in incredibly simple, and presumably independent, channels.

Next, how can we explain the seeming disconnect between the parathreshold accuracy findings of Green (1958) and others and our own low-capacity results with RTs? Actually, these results agree with a growing number of other investigations where performance is unlimited

or even super capacity at low signal-to-noise ratios, but they are limited to extremely limited at high levels of accuracy. In particular, the study of speech perception in the presence of lip reading has found phenomenal benefits, both behavioral (e.g., see Altieri, Pisoni, & Townsend, 2011; Altieri, Townsend, & Wenger, 2014; Sumbly & Pollack, 1954; Summerfield, 1987) and physiological (e.g., Stevenson & James, 2009), of the visual information when the signal-to-noise ratio is low, but depressed capacity at high echelons of accuracy. Lentz et al. (2014) also demonstrated this trend for detection of binaural tones added to binaural noise.

The apparent transition from a highly efficacious integration system such as associated with Green's integration model to a limited-capacity, parallel-decision model in our data also call for some type of phase-transition across levels of performance. We propose that a gain-controlled mechanism that encourages integration across channels when the sensory environment is degraded but tends to shut down such facilitation and integration when signal strength is elevated. This kind of meta-dynamic is only one aspect of many that beg for a unified treatment of audition that brings together RT and accuracy-based measures. To shed light on both the gain control mechanisms and the role of dynamics in accuracy experiments, we therefore have been working toward a more integrative theory that relies on both RT as well as accuracy measures. Townsend and Altieri (2012) have developed a capacity measure that is a function of both RTs and accuracy that can help in assessing human performance data as well as systems predictions. Donkin, Little, and Houpt (2014) used this measure and demonstrated that capacity varies based on whether subjects respond as accurately as possible or as quickly as possible, highlighting the importance of unifying accuracy and RT. We plan a series of studies founded on such unified strategies.

Acknowledgments

We would like to thank Amanda Hornbach and Rutendo Chikuku for assistance with data collection and analysis. Work was supported, in part, by grant No. R21DC013171 from NIH to the first author.

Notes

1. For two tones to be independently represented by the auditory system, one must consider the limitations of frequency selectivity and interactions that may occur between the two tones yielding unwanted perceptual changes. Generally speaking, tones should be "well-separated" in frequency so that they do not overlap in the neural representations (usually one-half octave is sufficient). They should also not have a common frequency that could be perceptible as an additional pitch.

2. Data presented here represent trials from a monaural presentation, but the full experiment included trials from a binaural presentation. Data from the binaural trials will be reported elsewhere.
3. For all observers and conditions, accuracy was very high, with error rates never falling below 3%.

References

- Altieri, N., Pisoni, D. B., & Townsend, J. T. (2011). Some behavioral and neurobiological constraints on theories of audiovisual speech integration: A review and suggestions for new directions. *Seeing and Perceiving, 24*, 513–539.
- Altieri, N., Townsend, J. T., & Wenger, M. J. (2014). A measure for assessing the effects of audiovisual speech integration. *Behavior Research Methods, 46*, 406–415.
- Buus, S., Schorer, E., & Zwicker, E. (1986). Decision rules in detection of simple and complex tones. *The Journal of the Acoustical Society of America, 80*, 1646–1657.
- Chocholle, R. (1944). Etude de la psychophysiologie de l'audition par la méthode des temps de réaction. *L'année psychologique, 45*(1), 90–131.
- Donders, F. C. (1869). On the speed of mental processes. *Acta Psychologica, 30*, 412–431.
- Donkin, C., Little, D., & Houpt, J. W. (2014). Assessing the speed-accuracy trade-off effect on the capacity of information processing. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 1183–1202.
- Durlach, N. I., Braida, L. D., & Ito, Y. (1986). Towards a model for discrimination of broadband signals. *The Journal of the Acoustical Society of America, 80*, 63–72.
- Egan, J. P., & Hake, H. W. (1950). On the masking pattern of a simple auditory stimulus. *The Journal of the Acoustical Society of America, 22*, 622–630.
- Egeth, H. E. (1966). Parallel vs serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics, 1*, 245–252.
- Eidels, A., Houpt, J. W., Altieri, N., Pei, L., & Townsend, J. T. (2011). Nice guys finish fast and bad guys finish last: Facilitatory vs inhibitory interaction in parallel systems. *Journal of Mathematical Psychology, 55*, 176–190.
- Eidels, A., Townsend, J. T., Hughes, H. C., & Perry, L. A. (2015). Evaluating perceptual integration: Uniting response-time and accuracy-based methodologies. *Attention, Perception, & Psychophysics, 77*(2), 659–680.
- Fletcher, H., & Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *Bell System Technical Journal, 12*, 377–430.
- Florentine, M., & Buus, S. (1981). An excitation-pattern model for intensity discrimination. *The Journal of the Acoustical Society of America, 70*, 1646–1654.
- Green, D. M. (1958). Detection of multiple component signals in noise. *The Journal of the Acoustical Society of America, 30*(10), 904–911.
- Green, D. M., McKey, M. J., & Licklider, J. C. (1959). Detection of a pulsed sinusoid in noise as a function of frequency. *The Journal of the Acoustical Society of America, 31*(11), 1446–1452.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford, England: John Wiley.
- Grice, G. R., Canham, L., & Gwynne, J. W. (1984). Absence of a redundant-signals effect in a RT task with divided attention. *Perception & Psychophysics, 36*, 565–570.

- Hicks, M. L., & Buus, S. (2000). Efficient across-frequency integration: Evidence from psycho-metric functions. *The Journal of the Acoustical Society of America*, 107, 3333–3342.
- Haupt, J. W., & Townsend, J. T. (2010). The statistical properties of the survivor interaction contrast. *Journal of Mathematical Psychology*, 54, 446–453.
- Haupt, J. W., & Townsend, J. T. (2011). An extension of SIC predictions to the Wiener coactive model. *Journal of Mathematical Psychology*, 55, 267–270.
- Haupt, J. W., & Townsend, J. T. (2012). Statistical measures for workload capacity analysis. *Journal of Mathematical Psychology*, 56(5), 341–355.
- Kohfeld, D. L. (1971). Simple RT as a function of stimulus intensity in decibels of light and sound. *Journal of Experimental Psychology*, 88, 251–257.
- Lentz, J. J., He, Y., & Townsend, J. T. (2014). A new perspective on binaural integration using response time methodology: Super capacity revealed in conditions of binaural masking release. *Frontiers in Human Neuroscience*, 8(641).
- Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 103–189). New York: Wiley.
- Marill, T. (1956). *Detection theory and psychophysics* (Unpublished doctoral dissertation). Massachusetts Institute of Technology. Dept. of Economics and Social Science.
- Marshall, L., & Brandt, J. F. (1980). The relationship between loudness and RT in normal hearing listeners. *Acta Oto-Laryngologica*, 90, 244–249.
- Neufeld, R. W. J., Townsend, J. T., & Jetté, J. (2007). Quantitative response time technology for measuring cognitive-processing capacity in clinical studies. In R. W. J. Neufeld (Ed.), *Advances in clinical cognitive science: Formal modeling and assessment of processes and symptoms* (pp. 207–238). Washington, DC: American Psychological Association.
- Schafer, T. H., & Gales, R. S. (1949). Auditory masking of multiple tones by random noise. *The Journal of the Acoustical Society of America*, 21(4), 392–397.
- Schlittenlacher, J., Ellermeier, W., & Arseneau, J. (2014). Binaural loudness gain measured by simple reaction time. *Attention, Perception, & Psychophysics*, 76(5), 1465–1472.
- Schröter, H., Ulrich, R., & Miller, J. (2007). Effects of redundant auditory stimuli on RT. *Psychonomic Bulletin & Review*, 14, 39–44.
- Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology*, 18, 105–139.
- Schweickert, R., & Townsend, J. T. (1989). A trichotomy: Interactions of factors prolonging sequential and concurrent mental processes in stochastic discrete mental (PERT) networks. *Journal of Mathematical Psychology*, 33, 328–347.
- Simon, J. R. (1967). Ear preference in a simple reaction-time task. *Journal of Experimental Psychology*, 75, 49–55.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153, 652–654.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 30, 276–315.
- Stevenson, R. A., & James, T. W. (2009). Neuronal convergence and inverse effectiveness with audiovisual integration of speech and tools in human superior temporal sulcus: Evidence from BOLD fMRI. *NeuroImage*, 44, 1210–1223.
- Sumbly, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, 26, 212–215.

- Summerfield, Q. (1987). Some preliminaries to a comprehensive account of audiovisual speech perception. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 3–51). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tanner Jr., P., W. (1956). Theory of recognition. *Journal of the Acoustical Society of America*, 28, 882–888.
- Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology*, 28, 363–400.
- Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, 108, 551–567.
- Townsend, J. T., & Altieri, N. (2012). An accuracy-response time capacity assessment function that measures performance against standard parallel predictions. *Psychological Review*, 119, 500–516.
- Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. N. J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 199–239). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Townsend, J. T., & Ashby, F. G. (1983). *The Stochastic Modeling of Elementary Psychological Processes*. Cambridge, UK: Cambridge University Press.
- Townsend, J. T., & Eidels, A. (2011). Workload capacity spaces: A unified methodology for response time measures of efficiency as workload is varied. *Psychonomic Bulletin & Review*, 18, 659–681.
- Townsend, J. T., & Landon, D. E. (1983). Mathematical models of recognition and confusion in psychology. *International Journal of Mathematical Social Sciences*, 4, 25–71.
- Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, 39, 321–359.
- Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, 33, 309–327.
- Townsend, J. T., & Wenger, M. J. (2004a). The serial-parallel dilemma: A case study in a linkage of theory and method. *Psychonomic Bulletin & Review*, 11, 391–418.
- Townsend, J. T., & Wenger, M. J. (2004b). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review*, 111, 1003–1035.
- Townsend, J. T., Yang, H., & Burns, D. M. (2011). Experimental discrimination of the world's simplest and most antipodal models: The parallel-serial issue. In H. Colonius & E. Dzhafarov (Eds.), *Descriptive and normative approaches to human behavior in the advanced series on mathematical psychology* (pp. 271–302). Singapore: World Scientific.
- van den Brink, W. A., & Houtgast, T. (1990a). Efficient across-frequency integration in short-signal detection. *The Journal of the Acoustical Society of America*, 87, 284–291.
- van den Brink, W. A., & Houtgast, T. (1990b). Spectro-temporal integration in signal detection. *The Journal of the Acoustical Society of America*, 88, 1703–1711.
- van Zandt, T., & Townsend, J. T. (2013). Designs for and analyses of response time experiments. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods, volume 1: Foundations* (pp. 260–285). New York: Oxford University Press.

- Wagner, E., Florentine, M., Buus, S., & McCormack, J. (2004). Spectral loudness summation and simple RT. *The Journal of the Acoustical Society of America*, 116, 1681–1686.
- Wenger, M. J., & Townsend, J. T. (2000a). Basic tools for attention and general processing capacity in perception and cognition. *Journal of General Psychology: Visual Attention*, 127, 67–99.
- Wenger, M. J., & Townsend, J. T. (2000b). Spatial frequencies in short-term memory for faces: A test of three frequency-dependent hypotheses. *Memory & Cognition*, 28, 125–142.

8 Characterizing and Quantifying Human Bandwidth

On the Utility and Criticality of the Construct of Capacity

Michael J. Wenger and Stephanie E. Rhoten

This is a chapter about optimism. Townsend's work has been credited with many things, but we know of no authors (other than ourselves) who have credited that work with the generation and preservation of—in the face of daunting conceptual and methodological challenges—optimism. We have commented on this with respect to the well-known problems associated with distinguishing serial from parallel processing (Townsend & Wenger, 2004a), noting that “the most widely held interpretation of the model-mimicking problem . . . is that it is a problem without a solution” (p. 418). Nothing, in fact, could be further from the truth. This is because, although Townsend provided the first and definitive characterization of the model-mimicking problem (e.g., Townsend, 1971, 1972; Townsend & Ashby, 1983), what is often overlooked is that, in these works and many that were to follow (including and not limited to Townsend, 1974, 1990; Townsend & Ashby, 1983; Townsend, Hu, & Evans, 1984; Townsend & Piotrowsky, 1981; Townsend & Schweickert, 1989), there were numerous suggestions for how the problem could be addressed, culminating in the comprehensive accomplishment of Townsend and Nozawa (1995).

The question of architecture, as emphasized repeatedly (e.g., Townsend & Wenger, 2004a, 2004b), cannot be addressed without also considering stopping rule, independence in rate, and capacity. The last of these constructs is our focus in this chapter. Townsend's work on capacity emerged during a period in which even the best of definitions were somewhat ad hoc, often circular, with little underlying theory and very little generality. In contrast, Townsend was progressively working toward a theory-based, general, non-parametric characterization of capacity, retaining the optimism of being able to find a solution when many investigators had tacitly abandoned hope.

We consider three periods of Townsend's work on capacity. The first is prior to 1995. This is the period during which Townsend was refining the theoretical characterization of the construct of capacity, while also developing quantitative tools to connect the theory with data, culminating in the paper with Nozawa (Townsend & Nozawa, 1995). The second covers the period from 1995 to 2004, ending with the publication of two papers (Townsend & Wenger, 2004a, 2004b) that extended the general characterization in