Why hearing impairment may degrade selective attention

Barbara G. Shinn-Cunningham[1,2]

[1] *Hearing Research Center, Departments of Cognitive and Neural Systems and Biomedical Engineering, Boston University, Boston, MA, 02215, USA*

[2] *Speech and Hearing Bioscience and Technology Program, Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, 02139, USA*

In everyday settings, the ability to selectively attend is critical for communication. Most normal-hearing listeners are able to selectively attend to a talker of interest in a sea of competing sources, and to rapidly shift attention as the need arises. However, hearing-impaired (HI) listeners and cochlear implant users have difficulty communicating when there are multiple sources. This paper reviews some of the processes governing selective attention in normal listeners. Results suggest that selective attention operates to select out perceptual "objects," and thus depends directly on the ability to separate a source of interest from a mixture of competing sources. In turn, this view suggests that one factor affecting how well HI listeners can communicate in everyday settings is their ability to perceptually organize the auditory scene.

INTRODUCTION

At any good cocktail party, the loud sounds of clinking glasses and exuberant voices add acoustically before entering your ears. In order to appreciate your companion's anecdote, you must filter out the extraneous sources and focus on his/her voice. At the same time, the sounds that you tune out are critical for maintaining awareness of your environment. Indeed, a sound that is a source of interference at one moment (the pompous man on your right) may become the very source you want to understand (e.g., when you realize he is discussing your boss).

Normal-hearing listeners are relatively good at listening selectively to a source of interest despite other sound sources in a crowded setting, and at switching attention as the need arises. However, these abilities are fragile. Even modest hearing impairments cause great difficulties when there are competing sources (e.g., Gatehouse and Akeroyd, 2006; Noble and Gatehouse, 2006). This often causes social isolation, as listeners opt out of even trying to participate rather than facing frustration and failure, especially because the cost of misunderstanding in a volatile social setting is often embarrassment or humiliation (e.g., Kochkin, 2005; Gatehouse and Akeroyd, 2006; Noble and Gatehouse, 2006; Edwards, 2007).

Understanding how normal-hearing listeners cope in complex settings is important for determining how to help impaired listeners. This paper describes the interference that arises in everyday settings, and the processes that we believe allow normal-hearing listeners to cope in these settings, focusing on the relationship between attention and object formation. We then consider why HI listeners may not be able to rely on these processes.

Auditory signal processing in hearing-impaired listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.). ISBN: 87-990013-1-4. Print: Centertryk A/S.

PERCEPTUAL INTERFERENCE IN EVERYDAY SETTINGS

Peripheral or "energetic" masking

Perceptual interference from competing sources may result simply because a masking source renders portions of the target inaudible. This kind of interference typically comes about because the masker overlaps (or nearly overlaps) in time and frequency with the target. As a result, some or all of the neural response to the target (e.g., at the level of the auditory nerve) will be distorted or even missing, because the sensory system is encoding a conflicting response to the masker. This kind of "energetic" masking has been the focus of study in pyschoacoustics for decades, and is relatively well understood.

Many natural sounds such as speech are spectro-temporally sparse, so energetic masking often affects only isolated points in time and frequency (Cooke, 2006). As a result, energetic masking is often not the factor limiting performance. Indeed, listeners are adept at using the good "glimpses" of a natural target sound to perceptually fill in inaudible portions of the target sound and make sense of a noisy, interrupted signal (Miller and Licklider, 1950; Warren, 1970).

Central or "informational" masking

Even when a target source is well represented on the auditory nerve, competing sources can still cause perceptual interference, referred to as "informational" masking (e.g., Durlach *et al.*, 2003). Typically, informational masking is defined as "all masking that is not energetic," a description that is, well, not particularly descriptive. Moreover, the use of a single term is misleading, as informational masking can occur at different levels of processing in the auditory system.

Informational masking has been linked with source segregation and attention (e.g., Leek *et al.*, 1991; Kidd *et al.*, 1994; Darwin and Hukin, 2000), a view we share. Informational masking has also been tied with stimulus similarity (i.e., similarity between target and maskers) and with stimulus uncertainty (e.g., randomness in the masker and/or target; e.g., see Kidd *et al.*, 2002a; Durlach *et al.*, 2003). While similarity and uncertainty do affect informational masking, we believe that explanations focused on stimulus attributes have little explanatory power. In particular, we believe that stimulus similarity and uncertainty directly affect source segregation and the ability to selectively attend (ideas developed further below).

By its very definition, informational masking occurs when information available at the auditory periphery is processed sub-optimally. Often, such sub-optimal processing reflects failures of attention. As discussed in detail below, attention acts on auditory objects, so attention and source segregation are intricately linked.

AUDITORY OBJECTS

(Lack of) definition of an auditory object

While the concept of an "object" is widely accepted in the field of vision, it is more controversial in audition (however, see the seminal work reviewed in Bregman, 1990). In particular, it is difficult to give a clear, precise definition of what an auditory "object" is or the rules that govern how an auditory object is formed. This difficulty arises in part because auditory objects are fluid and change over time (e.g., see Carlyon *et al.*, 2001), with few hard and fast rules governing their formation. Audible sound in a mixture is not allocated logically between the objects in the scene, and can contribute to multiple objects (Darwin, 1995; Whalen and Liberman, 1996; McAdams *et al.*, 1998) or to no object (Shinn-Cunningham *et al.*, 2007), depending on the situation. The state of the listener and his/her goals alter perceptual organization (e.g., see Darwin and Hukin, 1998; Cusack *et al.*, 2004; Sussman *et al.*, 2007). Expectations about a scene's content, the level of analysis that a listener is undertaking (listening to a symphony compared to listening to the English horn), and other cognitive factors can shift what constitutes an object at a given moment in time. Particularly for ambiguously structured stimuli, the perceptual organization of a scene can evolve over time and/or be bistable (e.g., see Pressnitzer and Hupe, 2006).

Yet, despite this lack of clear definition, most listeners have a solid intuitive feel for what constitutes an auditory object. In the cocktail party, the listener may perceive the woman to the left, the doorbell, the shattering plate, or the collegial slap on the back of the proud new father. Each of these objects can be thought of as an internal, perceptual estimate of the sound arising from a discrete physical source. Throughout this paper, the word "object" refers to a perceptual entity that, correctly or not, is perceived as coming from a discrete external, physical source.

Object formation

Auditory object formation occurs over different analysis scales. For sound elements that have contiguous spectral structure, formation relies primarily on this local spectro-temporal structure (Bregman, 1990; Darwin and Carlyon, 1995). The dominant (local) spectro-temporal cues for object formation include common onsets and offsets (more generally, common amplitude modulation), harmonic structure, and smooth changes in frequency over time. Thus, at the level of the speech syllable, sound is grouped primarily according to the features that also convey the signal's meaning. At this level of perceptual organization, location cues have a relatively weak influence (Darwin and Carlyon, 1995; Carlyon, 2004), but can have a measurable effect when other grouping cues are ambiguous (Darwin and Hukin, 1997; Shinn-Cunningham *et al.*, 2007).

Short-term objects such as syllables are streamed (linked together over time) through continuity and similarity of higher-order features such as location, pitch, timbre, sound quality, and even learned structure (word identity, grammatical structure, semantics; see Bregman, 1990; Darwin and Carlyon, 1995). At this level of organization, perceived location (but not necessarily underlying spatial cues such as interaural time differences or interaural level differences) has a strong influence on perceptual organization (Darwin and Hukin, 1999; Darwin and Hukin, 2000).

In summary, the influence of a particular cue or feature on object formation depends on the scale of the analysis. However, there is ample evidence that the ultimate perceptual organization of the scene, at all scales, depends on the preponderance of all evidence, including top-down influences (e.g., instructions to the listener, the expectations of the listener, etc.).

ATTENTION AND OBJECT FORMATION

Perception of foreground / background and the ubiquitous nature of objects

In all sensory modalities, the normal mode of analyzing a complex scene is to focus on one object (the foreground object) while others objects are in the perceptual background (Shomstein and Yantis, 2004; Duncan, 2006). In the vision literature, this mode of perceiving is described as coming about through a "biased competition" between perceptual objects (Desimone and Duncan, 1995). Biased competition takes place automatically and ubiquitously when there are multiple objects in a scene. The "winning" object depends both on its inherent salience (e.g., intensity) compared to other objects in the scene and the influence of volitional, top-down attention, which biases the competition to favor objects with desired features or attributes (Desimone and Duncan, 1995; Peers *et al.*, 2005; Yantis, 2005; Buschman and Miller, 2007).

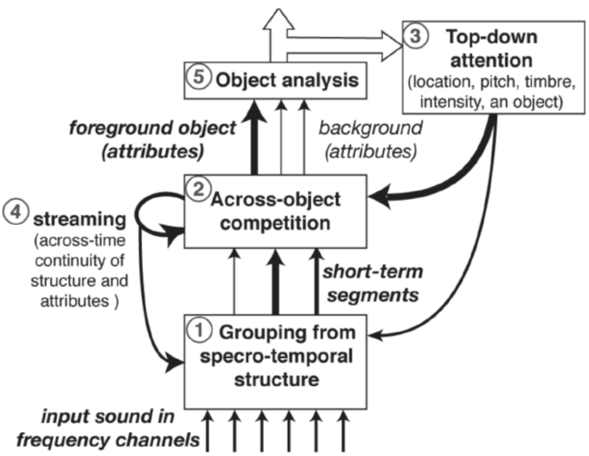

Fig. 1: Conceptual model relating object formation and attention.

Even when observers select what to attend (what to bring into the perceptual foreground) based on visual features, attention operates on perceptual objects (Desimone and Duncan, 1995; Serences *et al.*, 2005a). For instance, when attention is spatially focused, observers' sensitivity to other features that are part of the object at the attended location is also enhanced (e.g., Pestilli and Carrasco, 2005; Carrasco *et al.*, 2006). This shows that object formation is intricately linked with selective attention, and that the perceptual unit of attention is the "object." Thus, even though it is difficult to define what an object is, one cannot discuss perception in complex settings where selective attention is necessary without considering objects and object formation.

Most work on attention and objects is in the visual literature (for a recent review, see Knudsen, 2007), but we believe that similar principles govern auditory perception. There are similarities in the way that attention influences audition and vision (Busse *et al.*, 2005; Sanabria *et al.*, 2005; Serences *et al.*, 2005b; Turatto *et al.*, 2005; Shomstein and Yantis, 2006). Evidence suggests that attention acts on auditory objects, much as it enhances visual objects (Alain and Arnott, 2000; Scholl, 2001; Shinn-Cunningham *et* *al.*, 2005; Best *et al.*, 2007a). Moreover, listeners appear to attend actively to one and only one auditory object at a time (Cusack *et al.*, 2000; Vogel and Luck, 2002; Best *et al.*, 2006), consistent with the biased-competition model of visual attention.

Fig. 1 shows our conceptual model of these interactions in the auditory domain. 1) Short-term segments initially form based on local spectro-temporal grouping cues such as common onsets and offsets, harmonicity, and comodulation (Bregman, 1990; Darwin and Carlyon, 1995). 2) Competition first arises between short-term segments. Some segments may be inherently more salient than others (e.g., because of their intensity or distinctiveness; Conway *et al.*, 2001; Cusack and Carlyon, 2003), which biases the inter-segment competition (note the different arrow strengths leading from box 1 to box 2 in Fig. 1). 3) Top-down attention and 4) streaming (across-time source continuity) help modulate the competition, biasing it to favor objects with desirable features (top-down attention) and to continue to maintain attention on the object already in the foreground as it evolves over time (streaming) (Shinn-Cunningham *et al.*, 2005; Tata and Ward, 2005; Ahveninen *et al.*, 2006; Best *et al.*, 2006; Shomstein and Yantis, 2006; Winkowski and Knudsen, 2006). 5) As a result, one object is emphasized at the expense of others in the scene, which enhances analysis of this foreground object (Best *et al.*, 2005; Best *et al.*, 2006). This model suggests that proper grouping of the acoustic scene is necessary for listeners to selectively attend to a desired source.

Neural mechanisms of object formation and selection

There is no consensus about the neural mechanisms that support object formation or bring an object into the perceptual foreground. We argue that these processes are linked. Some of the spectro-temporal cues that support object formation undoubtedly are extracted automatically in early, low-level neural structures. For instance, extraction of auditory features important for object formation such as wideband onsets, comodulation, and harmonicity begins as early as the brainstem (e.g., see Langner, 1997; Arnott *et al.*, 2004; Ernst and Verhey, 2006). These cues, in turn, begin the process of object competition, which is modulated by top-down attention.

There are hints that neural synchrony is linked to object formation and selection (Fries *et al.*, 2001; Tiesinga, 2005; Womelsdorf *et al.*, 2006). Synchrony may initially be in response to a particularly salient source simply because neural responses to that source will be correlated, resulting in a highly synchronous response across a population of neurons. This spontaneous synchrony may be more effective at driving higher centers than responses to less salient sources, which will result in the salient source winning the biased competition for attention. The object that begins to emerge in the attentional focus can then cause modulatory feedback favoring its own attributes, resulting in a stronger neural bias to the winning object. Top-down attention can act to bias the initial responses to favor a source with a particular attribute, which will result in stronger, synchronous responses to the favored object. Results from visual studies suggest that just this kind of bootstrapping process works to drive the biased competition for visual objects; for instance, synchrony in neural responses is greater for attended objects than unattended objects (Tiesinga, 2005). Similar processes are likely responsible for realizing biased competition for auditory objects, as well.

RELEASE FROM INFORMATIONAL MASKING: SELECTIVE ATTENTION

We believe that normal-hearing listeners obtain release from informational masking by directing top-down attention to whatever perceptual object they are currently interested in analyzing (e.g., Shinn-Cunningham and Ihlefeld, 2004; Shinn-Cunningham, 2005; Best *et al.*, 2007a). By its very nature, the process of directing selective attention to a target depends on the proper perceptual formation of the target object. Normal-hearing listeners suffer most from informational masking when they cannot successfully focus attention on the target object. Specifically, failures in selective attention can come about from failures in 1) separating the target from the other sources (failures in object formation) and 2) directing attention to the correct object in the scene (failures in object selection).

Failures of object formation

As described above, object formation can be thought of at the syllable level (e.g., grouping together elements of sound that are logically contiguous in time and frequency), and at longer time scales (e.g., streaming together syllables over tens to hundreds of milliseconds).

Failures in object formation on the syllable level come about when the spectro-temporal features of an object are insufficient to separate the object from the other sources in a mixture (e.g., see Kidd *et al.*, 2002a; Best *et al.*, 2007b). This can occur for a variety of reasons:

1) the sound mixture may energetically mask so much of the target source that it cannot be segregated out from the background,
2) the mixture may contain other sources that have similar spectro-temporal structure and that tend to group with the desired source (see Fig. 2, left), or
3) the target signal itself may not be structured enough to support object formation, for instance, if the mixture contains ambiguous or conflicting cues (see Fig. 2, middle).

Failures to stream local objects across time can come about when there are multiple sources with similar higher-order features, such as when a listener hears a mixture of multiple male voices or the target is a set of tones that are masked by similar tones (e.g., Brungart, 2001; Kidd *et al.*, 2003). These failures can result in a target stream of speech that is corrupted by sound elements from a masker or from missing key elements or syllables.

Fig. 2 shows, by visual analogy, the kind of perceptual problems that can arise when object formation breaks down. In vision, spatial boundaries, texture, color, and similar features influence object formation. On the left, the general similarity of the features and elements of the image make it difficult to segregate words. As a result, viewers perceive the mixture as a connected mass. When this occurs, it takes time and extra cognitive effort to make sense of the words in the image. If color is used to differentiate the letters in the image, like-colored letters tend to group; however, if the letters making up the target word (focus on the middle of middle panel) fail to group together (even if they are perceptually separated from all interfering sources), analyzing the target word still requires extra effort. Finally, when the letters making up each word group together, each word can be attended and processed more automatically, resulting in less perceptual interference across words and understanding that requires less effort.

Fig. 2: Visual analogies of failed object formation. Left: target and competitors naturally group together to form a large object fails to represent any of the individual words. Middle: the target is not perceived as one unified object (direct attention to the middle of the image). Right; Understanding is clear when color helps letters group properly to form competing words.

Failures of object selection

Consistent with the theory of biased competition, volitional selection of an auditory object occurs through top-down attention. If the target object has features or properties that differentiate it from the other objects in a scene and if the listener knows this target-object feature a priori, they can direct attention to select the target out of the mixture of sounds they hear.

Failures in object selection can occur because a listener directs attention to the wrong object (either because they do not know what feature to attend, or because the other masking objects in a mixture are similar to the target object in a particular feature; e.g., Kidd *et al.*, 2005). However, sometimes, even when the listener is sure of what object they wish to attend, they may fail. This second kind of failure typically occurs because some competing object is inherently more salient than the target object (e.g., are much louder than the target object), or because the target and maskers are not sufficiently distinct to ensure proper target selection (Brungart, 2001; Darwin *et al.*, 2003; Kidd *et al.*, 2005). In short, sometimes, the top-down bias of selective attention is not sufficient to override other cues and win the biased competition.

Fig. 3 illustrates the influence of bottom-up salience on attention, again by visual analogy. In this example, there are a number of discrete words that form into discrete objects based primarily on the spatial proximity of the letters within, compared to across, the words of the image. Thus, object formation is not an issue; letters are properly associated together to form discrete, meaningful words, and the observer can analyze each at will.

Fig. 3: Visual analogy illustrating how object selection can be driven by bottom-up salience.

For this image, the phrase “bottom-up” pops out of the image because it is darker than all of the other words. Because it is different from and more salient than the other words, attention is automatically drawn to this phrase even in the absence of any top-down desire to attend to it. However, if a viewer is specifically told to look at the bottom left corner of the image, the phrase “top-down” becomes the focus of attention. This shows how top-down object selection can override bottom-up salience to select the focus of attention. However, in order to choose the correct word from the mixture, the observer must be told some feature (here, spatial location) that differentiates the target from the other words in the image.

This image illustrates that biased competition for attention is affected both by bottom-up and top-down pressures. Depending on the strength of the bottom-up cues for attention, top-down attention may be insufficient to ensure proper object selection. Moreover, the more unique and distinct the target features are, the more precise and effective top-down attention is in enhancing the target and suppressing the maskers. Thus, object selection is a probabilistic competition that depends on interactions between bottom-up and top-down biases.

Understanding stimulus similarity and stimulus uncertainty

As noted above, stimulus similarity and stimulus uncertainty affect informational masking (see Kidd *et al.*, 2002a; Durlach *et al.*, 2003). However, we argue that these stimulus properties affect performance by interfering with object formation and object selection.

Similarity between target and masker can cause either or both of the processes of object formation and object selection to fail. Similarity can cause the target and masker to be perceived as part of the same, larger perceptual object, which will result in poorer sensitivity to the content of the target. Even if target and masker are perceptually segregated into distinct objects, similarity of these objects may make it difficult to direct attention to the correct object. Uncertainty also can interfere with object selection, either because the listener is unsure of how to direct top-down attention to select the target object, or because the inherent salience of new, novel events (e.g., randomly varying maskers) draws exogenous attention that is too strong to be overcome by top-down attention.

Thus, studies that focus on stimulus properties are consistent with the view that informational masking comes about through failures of selective attention. In trying to understand informational masking, it is useful to frame the problem in terms of object formation and object selection, rather than in terms of the similarity and uncertainty in the stimuli.

IMPLICATIONS FOR HEARING-IMPAIRED LISTENERS

Subjective reports from HI listeners suggest that they have difficulty in situations where normal-hearing listeners rely on selective attention. Discouraged listeners make comments like "My hearing aid amplifies the background as much as the speech" (Kochkin, 2005; Gatehouse and Akeroyd, 2006; Noble and Gatehouse, 2006; Edwards, 2007). Reports suggest that HI listeners have most difficulty when attention must shift rapidly from source to source, like at a cocktail party, resulting in social isolation (Gatehouse and Akeroyd, 2006; Noble and Gatehouse, 2006).

Poor peripheral auditory representation

Impaired listeners have reduced temporal and spectral acuity compared to normal-hearing listeners (e.g., Leek and Summers, 2001; Gatehouse *et al.*, 2003; Deeks and Carlyon, 2004; Bernstein and Oxenham, 2006; Carlyon *et al.*, 2007). If the features that convey speech meaning are degraded due to reduced audibility and diminished spectro-temporal resolution, speech intelligibility will be degraded even in quiet. Moreover, HI listeners may suffer from effective increases in the amount of energetic masking due to the reduced spectral selectivity of their peripheral auditory filters and higher-than-normal absolute thresholds. Together, these factors cause less of a target source to be audible to a HI listener than would be audible for a normal-hearing listener in the same acoustic setting. Most past work on how to ameliorate hearing impairment addresses these kinds of issues. For instance, many hearing-aid algorithms are designed to amplify sound and ensure that it falls within the dynamic range of hearing, and to reduce energetic interference from noise or background sounds.

Failures of object formation?

While energetic masking is undoubtedly a factor affecting how HI listeners cope in complex settings, HI listeners are also likely to have difficulty properly grouping sound sources. The very spectro-temporal cues that convey speech meaning also drive short-term grouping (Bregman, 1990; Darwin and Carlyon, 1995). A less-robust representation of spectro-temporal content in impaired listeners may cause problems with object formation. For instance, the onsets, offsets, modulation, and harmonic structure important for forming objects over short time scales (e.g., forming syllables from a sound mixture composed of many talkers) are less perceptually distinct for HI listeners than normal-hearing listeners (Leek and Summers, 2001; Mackersie *et al.*, 2001; Bacon and Opie, 2002; Kidd *et al.*, 2002b; Bernstein and Oxenham, 2006; Noble and Gatehouse, 2006). For the same reasons, robust location, pitch, and harmonic cues may not be available to HI listeners, further impairing their ability to properly separate the mixture into streams.

Failures of object selection?

As in the visual analogy of Fig. 2, if HI listeners fail to properly form auditory objects, they will have difficulty selectively attending to a target. When objects form properly, biased competition between objects works to suppress the objects outside the focus of attention. When objects fail to form properly, the competing sources will not be suppressed effectively, and therefore will cause greater perceptual interference. In addition, loss of spectro-temporal detail in the periphery may “muddy” perception of higher-order features that distinguish target from masker. For instance, impairments in pitch perception and sound localization may degrade how precisely HI listeners are able to focus attention on a target object, even if it is properly formed.

SUMMARY AND CONCLUSIONS

Normal-hearing listeners are able to direct top-down attention to select desired auditory objects from out of a sound mixture. Because perceptual objects are the basic units of attention, proper object formation is important for being able to selectively attend. Spectro-temporal structure of sound determines how objects form. However, spectro-temporal detail is not encoded robustly in HI listeners. Normal-hearing listeners can direct top-down attention to a desired object, enhancing it and suppressing competing maskers. In contrast, HI listeners may have difficulty in properly forming objects as well as selecting them. These difficulties are likely to contribute to the difficulties HI listeners have in settings with competing sound sources.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Deafness and Communication Disorders, the Air Force Office of Scientific Research, the Office of Naval Research, and the National Science Foundation. The ideas presented here were developed, challenged, and refined through extensive discussions with my friends and colleagues, especially Gin Best, Antje Ihlefeld, Erick Gallun, Chris Mason, Gerald Kidd, Steve Colburn, Nat Durlach, and Nicole Marrone.

REFERENCES

Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F. H., Sams, M., Shinn-Cunningham, B. G., Witzel, T., and Belliveau, J. W. (**2006**). "Task-modulated "what" and "where" pathways in human auditory cortex," Proc. Natl. Acad. Sci., **103**, 14608-13.

Alain, C. and Arnot, S. R. (**2000**). "Selectively attending to auditory objects," Front. Biosci., **5**, D202-12.

Arnott, R. H., Wallace, M. N. Shackleton, T. M., and Palmer, A. R. (**2004**). "Onset neurones in the anteroventral cochlear nucleus project to the dorsal cochlear nucleus," J. Assoc. Res. Otolaryngol., **5**, 153-70.

Bacon, S. P., and Opie, J. M. (**2002**). "Modulation detection interference in listeners with normal and impaired hearing," J. Speech. Lang. Hear. Res., **45**, 392-402.

Bernstein, J. G., and Oxenham, A. J. (**2006**). "The relationship between frequency selectivity and pitch discrimination: sensorineural hearing loss," J. Acoust. Soc. Am., **120**, 3929-45.

Best, V., Gallun, F. J., Carlile, S., and Shinn-Cunningham, B. G. (**2007b**). "Binaural interference and auditory grouping," J. Acoust. Soc. Am., **121**, 420-432.

Best, V., Gallun, F. J. Ihlefeld, A., and Shinn-Cunningham, B. G. (**2006**). "The influence of spatial separation on divided listening," J. Acoust. Soc. Am., **120**, 1506-1516.

Best, V., A. Ihlefeld, and Shinn-Cunningham, B. (**2005**). "The effect of spatial layout in a divided attention task," International Conference on Auditory Display, Limerick, Ireland.

Best, V., Ozmeral E. J,. and Shinn-Cunningham, B. G. (**2007a**). "Visually guided attention enhances target identification in a complex auditory scene," J. Assoc. Res. Otolaryngol., **8**, 294-304.

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

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

Buschman, T. J., and Miller, E. K. (**2007**). "Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices," Science, **315**, 1860-2.

Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., and Woldorff, M. G. (**2005**). "The spread of attention across modalities and space in a multisensory object," Proc. Natl. Acad. Sci., **102**, 18751-18756.

Carlyon, R. P. (**2004**). "How the brain separates sounds," Trends Cogn. Sci., 8, 465-471.

Carlyon, R. P., Cusack, R., Foxton, J. M., and Robertson, I. H. (**2001**). "Effects of attention and unilateral neglect on auditory stream segregation," J. Exp. Psychol. Hum. Percept. Perf., **27**, 115-127.

Carlyon, R. P., Long, C. J., Deeks, J. M., and McKay, C. M. (**2007**). "Concurrent sound segregation in electric and acoustic hearing," J. Assoc. Res. Otolaryngol., **8**, 119-133.

Carrasco, M., Giordano, A. M., and McElree, B. (**2006**). "Attention speeds processing across eccentricity: feature and conjunction searches," Vision Res., **46**, 2028-40.

Conway, A. R., Cowan, N., and Bunting, M. F. (**2001**). "The cocktail party phenomenon revisited: the importance of working memory capacity," Psychon. Bull. Rev., **8**, 331-335.

Cooke, M. (**2006**). "A glimpsing model of speech perception in noise," J. Acoust. Soc. Am., **119**, 1562-1573.

Cusack, R., and Carlyon, R. P. (**2003**). "Perceptual asymmetries in audition," J. Exp. Psychol. Hum. Percept. Perf., **29**, 713-725.

Cusack, R., Carlyon, R. P., and Robertson, I. H. (**2000**). "Neglect between but not within auditory objects," J. Cogn. Neurosci., **12**, 1056-1065.

Cusack, R., Deeks, J., Aikman, G., and Carlyon, R. P. (**2004**). "Effects of location, frequency region, and time course of selective attention on auditory scene analysis," J. Exp. Psychol. Hum. Percept. Perf., **30**, 643-656.

Darwin, C. J. (**1995**). "Perceiving vowels in the presence of another sound: A quantitative test of the "Old-plus-New" heuristic," in Levels in Speech Communication: Relations and Interactions: A Tribute to Max Wajskop, edited by J. C Sorin, H Meloni, and J Schoenigen, 1-12.

Darwin, C. J., Brungart, D. S., and Simpson, B. D. (**2003**). "Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers," J. Acoust. Soc. Am., **114**, 2913-2922.

Darwin, C. J., and Carlyon, R. P. (**1995**). "Auditory grouping," in Hearing, edited by B. C. J. Moore, 387-424.

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.

Darwin, C. J., and Hukin, R. W. (**1998**). "Perceptual segregation of a harmonic from a vowel by interaural time difference in conjunction with mistuning and onset asynchrony," J. Acoust. Soc. Am., **103**, 1080-1084.

Darwin, C. J., and Hukin, R. W. (**1999**). "Auditory objects of attention: the role of interaural time differences," J. Exp. Psychol. Hum. Percept. Perf., **25**, 617-629.

Darwin, C. J., and Hukin, R. W. (**2000**). "Effectiveness of spatial cues, prosody, and talker characteristics in selective attention," J. Acoust. Soc. Am., **107**, 970-977.

Deeks, J. M., and Carlyon, R. P. (**2004**). "Simulations of cochlear implant hearing using filtered harmonic complexes: implications for concurrent sound segregation," J. Acoust. Soc. Am., **115**, 1736-1746.

Desimone, R., and Duncan, J. (**1995**). "Neural mechanisms of selective visual attention," Annu. Rev. Neurosci., **18**, 193-222.

Duncan, J. (**2006**). "EPS Mid-Career Award 2004: brain mechanisms of attention," Q J Exp Psychol (Colchester), **59**, 2-27.

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

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., and Kidd, G., Jr. (**2003**). "Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity," J. Acoust. Soc. Am., **114**, 368-379.

Edwards, B. (**2007**). "The future of hearing aid technology," Trends Amplif, **11**, 31-45.

Ernst, S. M., and Verhey J. L. (**2006**). "Role of suppression and retro-cochlear processes in comodulation masking release," J. Acoust. Soc. Am., **120**, 3843-52.

Fries, P., Reynolds, J. H., Rorie, A. E., and Desimone, R. (**2001**). "Modulation of oscillatory neuronal synchronization by selective visual attention," Science, **291**, 1560-1563.

Gatehouse, S., and Akeroyd, M. (**2006**). "Two-eared listening in dynamic situations," Int. J. Audiol., **45** Suppl 1, S120-124.

Gatehouse, S., Naylor, G., and Elberling, C. (**2003**). "Benefits from hearing aids in relation to the interaction between the user and the environment," Int. J. Audiol., **42** Suppl. 1, S77-85.

Kidd, G., Jr., Arbogast, T. L., Mason, C. R., and Gallun, F. J. (**2005**). "The advantage of knowing where to listen," J. Acoust. Soc. Am., **118**, 3804-15.

Kidd, G., Jr., Arbogast, T. L., Mason, C. R., and Walsh, M. (**2002**). "Informational masking in listeners with sensorineural hearing loss," J. Assoc. Res. Otolaryngol., **3**, 107-119.

Kidd, G., Jr., Mason, C. R., and Arbogast, T. L. (**2002**). "Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns," J. Acoust. Soc. Am., **111**, 1367-1376.

Kidd, G., Jr., 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.

Kidd, G., Mason, C. R., Deliwala, P. S., Woods, W. S., and Colburn, H. S. (**1994**). "Reducing informational masking by sound segregation," J. Acoust. Soc. Am., **95**, 3475-3480.

Knudsen, E. I. (**2007**). "Fundamental components of attention," Annu. Rev. Neurosci., **30**, 57-78.

Kochkin, S. (**2005**). "MarkeTrak VII: Customer satisfaction with hearing instruments in the digital age," The Hearing Journal, **58**, 30-37.

Langner, G. (**1997**). "Neural processing and representation of periodicity pitch," Acta Otolaryngol Suppl, **532**, 68-76.

Leek, M. R., Brown, M. E., and Dorman, M. F. (**1991**). "Informational masking and auditory attention," Percept. Psychophys., **50**, 205-214.

Leek, M. R., and Summers, V. (**2001**). "Pitch strength and pitch dominance of iterated rippled noises in hearing-impaired listeners," J. Acoust. Soc. Am., **109**, 2944-2954.

Mackersie, C. L., Prida, T. L., and Stiles, D. (**2001**). "The role of sequential stream segregation and frequency selectivity in the perception of simultaneous sentences by listeners with sensorineural hearing loss," J. Speech. Lang. Hear. Res., **44**, 19-28.

McAdams, S., Botte, M. C., and Drake, C. (**1998**). "Auditory continuity and loudness computation," J. Acoust. Soc. Am., **103**, 1580-1591.

Miller, G. A., and Licklider, J. C. R. (**1950**). "The intelligibility of interrupted speech," J. Acoust. Soc. Am., **22**, 167-173.

Noble, W., and Gatehouse, S. (**2006**). "Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the Speech, Spatial, and Qualities of Hearing Scale (SSQ)," Int. J. Audiol., **45**, 172-181.

Peers, P. V., Ludwig, C. J., Rorden, C., Cusack, R., Bonfiglioli, C., Bundesen, C., Driver, J., Antoun, N., and Duncan, J. (2005). "Attentional functions of parietal and frontal cortex," Cereb Cortex, **15**, 1469-1484.

Pestilli, F., and Carrasco, M. (**2005**). "Attention enhances contrast sensitivity at cued and impairs it at uncued locations," Vision Res., **45**, 1867-1875.

Pressnitzer, D., and Hupe, J. M. (**2006**). "Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization," Curr. Biol., **16**, 1351-1357.

Sanabria, D., Soto-Faraco, S., Chan, J., and Spence, C. (**2005**). "Intramodal perceptual grouping modulates multisensory integration: evidence from the crossmodal dynamic capture task," Neurosci. Lett., **377**, 59-64.

Scholl, B. J. (**2001**). "Objects and attention: the state of the art," Cognition, **80**, 1-46.

Serences, J. T., Liu, T., and Yantis, S. (**2005b**). "Parietal mechanisms of switching and maintaining attention to locations, objects, and features," in Neurobiology of

Attention, edited by L. Itti, G. Rees and J. Tsotsos, 35-41.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E. and Yantis, S. (**2005a**). "Coordination of voluntary and stimulus-driven attentional control in human cortex," Psychol. Sci., **16**, 114-122.

Shinn-Cunningham, B. G. (**2005**). "Influences of spatial cues on grouping and understanding sound." Forum Acusticum 2005, Budapest, Hungary, CD.

Shinn-Cunningham, B. G., and Ihlefeld, A. (**2004**). "Selective and divided attention: Extracting information from simultaneous sound sources." International Conference on Auditory Display, Sydney, Australia

Shinn-Cunningham, B. G., Ihlefeld, A., Satyavarta, and Larson, E. (**2005**). "Bottom-up and top-down influences on spatial unmasking," Acta Acust., **91**, 967-979.

Shinn-Cunningham, B. G., Lee, A. K., and. Oxenham, A. J (**2007**). "A sound element gets lost in perceptual competition," Proc. Natl. Acad. Sci., 104.

Shomstein, S., and Yantis, S. (**2004**). "Configural and contextual prioritization in object-based attention," Psychon. Bull. Rev., **11**, 247-253.

Shomstein, S., and Yantis, S. (**2006**). "Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention," J. Neurosci., **26**, 435-439.

Sussman, E. S., Horvath, J., Winkler, I. and Orr, M. (**2007**). "The role of attention in the formation of auditory streams," Percept. Psychophys., **69**, 136-52.

Tata, M. S., and Ward, L. M. (**2005**). "Spatial attention modulates activity in a posterior "where" auditory pathway," Neuropsychologia, **43**, 509-516.

Tiesinga, P. H. (**2005**). "Stimulus competition by inhibitory interference," Neural Comput, **17**, 2421-2453.

Turatto, M., Mazza, V., and Umilta, C. (**2005**). "Crossmodal object-based attention: auditory objects affect visual processing," Cognition, **96**, B55-64.

Vogel, E. K., and Luck, S. J. (**2002**). "Delayed working memory consolidation during the attentional blink," Psychon. Bull. Rev., **9**, 739-743.

Warren, R. M. (**1970**). "Perceptual restoration of missing speech sounds," Science, **167**, 392-393.

Whalen, D. H., and Liberman, A. M. (**1996**). "Limits on phonetic integration in duplex perception," Percept. Psychophys., **58**, 857-870.

Winkowski, D. E., and Knudsen, E. I. (**2006**). "Top-down gain control of the auditory space map by gaze control circuitry in the barn owl," Nature, **439**, 336-339.

Womelsdorf, T., Fries, P., Mitra, P. P., and Desimone, R. (**2006**). "Gamma-band synchronization in visual cortex predicts speed of change detection," Nature, **439**, 733-736.

Yantis, S. (**2005**). "How visual salience wins the battle for awareness," Nat. Neurosci., **8**, 975-977.