INTERSPEECH 2016
September 8–12, 2016, San Francisco, USA

Release from energetic masking caused by repeated patterns of glimpsing windows

Maury Lander-Portnoy[1]

[1]Department of Linguistics
University of Southern California, USA
landerpo@usc.edu

Abstract

The study of auditory masking not only provides data for how healthy and impaired listeners perform in adverse listening conditions, and thereby approximates their ability to perceive speech in the noisy environments of everyday life, but also provides insights into the mechanisms that underly the detection and perception of speech. Previous studies, (Pollack 1955) (Festen & Plomp 1990) (Cooper et al. 2015), have manipulated noise maskers in an attempt to observe the relationship between modulation of the type or characteristics of masking noise to subjects ability to detect or recognize a target signal. In this experiment, long term average spectrum speech shaped noise maskers were modulated to allow either short or long glimpsing (Cooke 2005) windows, during which the target signal was unmasked, in one second long morse code patterns of eight windows. The results from 60 participants with normal hearing showed that subjects performed significantly better on trials of an open set word recognition task when the pattern of glimpsing windows repeated twice before presentation of the masked signal than a control with the same glimpsing windows during the signal but different beforehand and one with the same amount of noise masking in random patterns before and during the target.

Index Terms: speech perception, speech perception in noise, energetic masking

1. Introduction

1.1. Auditory masking

Dialogue occurs in a variety of environments throughout everyday life. While some of these auditory scenes provide a blank canvas for linguistic interchange, most contain at least some level of background noise. Masking, the obstructing of the detection or comprehension of a target signal, can be either asynchronous (the masker precedes or follows the target) or synchronous (the target is partially or wholly contained within the masker). While plausibly occurring in this experiment, the conditions were counterbalanced to control for any asynchronous masking. Thus, the focus here will be on synchronous maskers, the kind manipulated for the experiment presented in this paper.

Synchronous maskers are traditionally divided into two groups depending on the way in which they obstruct the perception of the target signal. The first type, informational masking, concerns the presentation of information similar and in close proximity to the target signal. This masker hinders the processing of the correct signal by offering distracting information that appears similar to the target and thus is often confused with the target in the processing of the input. While the target is still perceivable, the overlap of target and masker makes it difficult for cognitive processes to tease them apart. Top-down processing plays a crucial role in the release from informational masking. Because of its predication on top-down information, it has been hypothesized to be a more central cognitive process occurring further downstream in the auditory transduction pathway[1][2], and therefore the things that confound its operation are higher level processes such as attention or perceptual grouping[3]. Its counterpart, energetic masking, is conversely conceived of as a peripheral masking phenomenon[4][5].

Energetic masking is thought to hinder perception by obscuring the target signal with surrounding noise. While the characteristics of the signal and noise are quite different, as in the case of speech and white noise, too much noise in the input prevents proper signal processing. The difficulty experienced with this type of masker is thought to be due to an overlap of noise and target in the peripheral sensory organs. This means that areas of the system being used to detect the target are also used to detect the noise, and the signal becomes washed out by interference. Because of its hypothesized peripheral nature, we find energetic masking to utilize low level confounds such as exhibiting the same spectral characteristics as speech's long term average spectrum (LTAS) but with none of the temporal information included in its envelope[6]. The closer the peripheral activation by the noise is to the activation by the target, the more interference and blocking to sensory resources the masker can provide. A useful line of inquiry when studying masking is by what means we can negate its effects, called "release from masking". The release from masking is important to study as by learning what defeats masking, we can gain better insight into the process by which it obscures perception. Many mechanisms are studied with regards to release from masking but one of the main ones is auditory stream segregation.

1.2. Auditory scene analysis

Auditory scene analysis[7] is the synthesis of temporally and spectrally disparate acoustic information into cohesive auditory percepts (for a review see [8]). These entities are referred to as auditory objects and their formation is key in interacting with the auditory world. The question of what mechanisms are utilized in this process has led researchers to examine the criteria that are used in formation and separation of these auditory objects. Two important variables in auditory stream segregation are time and attention, that is, auditory stream segregation is an online process that takes time to occur and is a process that must be attended to [9][10]. It does not occur instantaneously, and if attention is shifted away from an auditory object, streaming rapidly resets and the process must start all over again[11]. It is this stream formation and segregation, or something similar, that we hypothesized might provide release from energetic

http://dx.doi.org/10.21437/Interspeech.2016-1571

masking if provided the necessary information to structure and sort the acoustic signal. In addition to auditory stream segregation, the last important topic that informs our experiment is that of speech perception in adverse listening conditions.

1.3. Speech perception in noise (SPiN)

Theoretical accounts of speech perception in adverse listening conditions, such as those introduced in our experiment, are very important for the design and interpretation of the experiment itself. A predominant theory in the field of SPiN is that of glimpsing. In the glimpsing model, people take advantage of periods of low power in the noise to "glimpse" the target signal beyond it [12][13]. This was originally tested by regularly interrupting speech with noise in a "Picket Fence" pattern[14]. This allowed participants to gain regular predictable glimpses of the target signal beneath the masking noise, improving their performance. Providing these windows to participants does not necessarily have to occur in a completely cyclical and predictable nature however. Rather than a "Picket Fence" pattern, the slats of the fence could be stretched or shrunk and they could be moved left and right to provide for different interesting patterns of windows in which to glimpse the target. This glimpsing window manipulation, coupled with the aforementioned ability to lock onto regular temporal and spectral information in auditory stream construction and segregation, provides the basis for the experiment presented here. The application of auditory stream segregation to energetic masking imbued with temporal and spectral regularities would provide a novel method of release from energetic masking. If this were the case, we would need to allow time for the streaming to build up and for subjects to pay attention to the noise to be streamed. Given a few cycles of glimpsing windows occurring in a repeating pattern, we hypothesize that participants may be able to learn the properties of this pattern and use this information online to aid in speech perception.

2. Methods

2.1. Participants

60 college-age undergraduates at the University of Southern California participated in the study in exchange for Psychology department course credit. All participants had self-reported normal hearing and were monolingual speakers of English. Only participants who completed all three tasks in their entirety were considered to provide valid data for analysis. In total this led to 6 participants' data being thrown out, meaning 66 subjects were run in total including those with incomplete data.

2.2. Procedure

Participants sat in a noise attenuating booth for the duration of each task. Stimuli were presented using the Paradigm software from a PC in mono at a sampling rate of 44.1 KHz at 16 bits over headphones at a comfortable level consistent across participants. Once the stimulus had finished, participants typed responses in a free response text box and then advanced to the next trial manually. Following the speech perception task, subjects participated in a Simon task of executive function. In Simon, a red or yellow square occurred variably on the left or right side of the screen. Depending on the square's color, and *not* its position on screen, participants should have clicked either the left or right mouse button. Trials in which the correct mouse button differed from the side of the screen of square presentation are thus considered incongruent, and require executive function to complete correctly. After a response was logged, the next trial was automatically initiated. Finally, subjects participated in a digit span recall and ordering task of working memory in which stimuli were again presented visually on the PC monitor and responses were input via keyboard into a free response text box. Subjects then used the keyboard to initiate the next trial.

2.3. Stimuli

Target items were multisyllabic words selected from The Nationwide Speech Project Corpus[15]. Target items were manually extracted from the corpus' sound files using Praat[16] to ensure the beginning of the target sound file coincided with the target word's onset. Stimuli were masked by speech shaped noise (SSN) correlated to the long term average spectrum (LTAS) of the aggregate of all 108 target words. While SSN correlated to the LTAS of each individual token would have provided more difficult energetic masking, by virtue of the two spectra being more closely overlapped, and thus obscuring the peripheral perception of the signal more, such an approach would also undesirably give participants spectral information about the target to be heard well before the presentation of the word. Because the masking noise was not correlated to the LTAS of that specific trial's target, the spectrum of the trial's target only roughly matched that of its masker. After filtering was finished, noise samples' power was then set relative to the power of that trial's target at a signal-to-noise ratio (SNR) of -10dB.

The noise samples were then divided into 8Hz windows for organization into 1-second-long patterns of long or short glimpsing windows. Short glimpsing windows allowed a noise-free portion of 20% with the remaining 80% maintaining its noise. Long glimpsing windows conversely allowed an 80% noise-free portion, leaving just 20% of the noise. The amount of long and short windows within the 8 window pattern was kept consistent and equal (4 longs and 4 shorts in each pattern). The chosen 8Hz window frequency fits nicely within the 2-12Hz range that is a popular frequency cited for several neurophysiological oscillation patterns correlating to speech perception [17][18]. We then introduced jitter into the patterns to allow for 4 unique window types. The jitter was introduced in the windows by centering the period of noiselessness with respect to the window. This means that the middle of the period of noiselessness and the middle of the window coincided. A visual representation to aid in conceptualization of the noise patterns and glimpsing windows can be found in Figure 1.

The different conditions utilized these 8 window patterns differently. Each trial was 3 seconds long and thus contained 3 patterns of windows: two presented before the target, in the preamble, and one masking the presentation of the target. Condition 1, in which we hypothesized streaming might take place, contained a pattern that repeated twice in the preamble (for reference we call this repeating pattern "Pattern A") as well as a third time during the masking window. Pattern A not only repeated 3 times within the trial, but was also reused for every Condition 1 trial. Because the noise pattern repeats twice before the presentation of the masker and three times in total, we predict that participants will be able to recognize and utilize this predictable temporal information online to aid in word recognition and thus perform significantly better in Condition 1 trials than any other condition. Additionally, participants' performance may increase over the course of the experiment as they gain more exposure to Pattern A or as they improve in their ability to utilize the repeating temporal information.

Condition 2, a control condition to test whether participants were learning Pattern A throughout the course of the experiment (rather than simply using the information in online processing and discarding it), contained two random novel patterns in the preamble, and Pattern A as the target masker. The two random patterns in the preamble would not only be unrelated to Pattern A, but would also have never been heard by the participant prior to that trial. Due to the three patterns in Condition 2 being different from each other, we predict that no online pattern recognition could occur, and therefore any improvement in this condition would be due to past exposure to Pattern A. While we predict participants will perform worse on Condition 2 trials than Condition 1 trials due to lack of online information, it is possible exposure to Pattern A may provide participants with enough across trial information to perform significantly better on Condition 2 trials than baseline trials.

Condition 3, a control condition to test participants' baselines throughout the experiment, contained two random patterns in the preamble and a random pattern as a masker. All three patterns would not only be unique from each other, but also from any patterns the participant had seen previously in the experiment. Because there was no repeating information present, either within trials, as in Condition 1, or across trials, as in both Condition 1 and Condition 2, we predict that participants will not be able to utilize any information gained prior to the presentation of the masked target. As such, this condition will serve as a baseline for a participant's performance on a SPiN task, and any improvement in this condition signifies improvement at the task which should occur in all conditions, rather than learning or online processing advantages. We therefore predict that participants will perform the worst in Condition 3. In summary, for the three conditions, with 36 trials per condition, we are left with the following format for trials:

Table 1: *Pattern structure of conditions*

	Preamble 1	Preamble 2	Masking Pattern
Cond. 1	Pattern A	Pattern A	Pattern A
Cond. 2	36 Randoms	36 Randoms	Pattern A
Cond. 3	36 Randoms	36 Randoms	36 Randoms

Condition 1

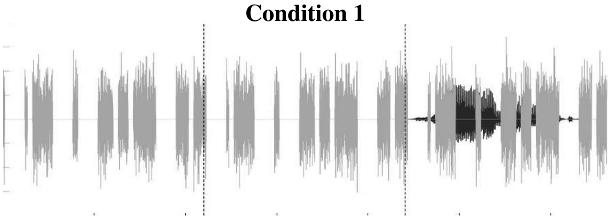

Figure 1: *Wave form demonstrating the morse code like pattern of the noise (light gray) and its masking of the target (dark gray). Dotted vertical lines have been placed to demarcate the boundary between patterns. In this example, condition 1, the pattern is Pattern A which repeats 3 times in the trial.*

3. Results

Participant performance on each trial of the SPiN task was initially automatically checked for exact matches (ignoring capitalization) to a correct response list by the Paradigm Software and given a score of 1 if correct and 0 if incorrect. The responses were further automatically checked for whitespace errors in MATLAB and finally hand checked for misspellings.

The following statistics can be found summarized visually in Figure 2. Mean accuracy in condition 1, in which we hypothesized streaming might take place, was 66%. The mean accuracy in condition 2, in which we hypothesized learning might take place, was 61%. The mean accuracy in condition 3, the random baseline condition, was also 61%.

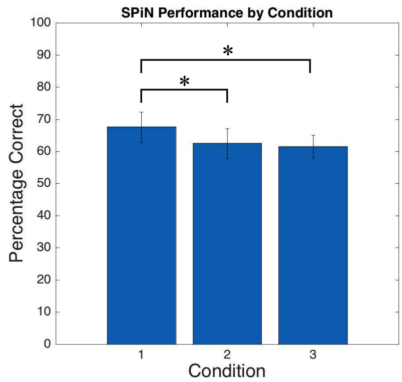

Figure 2: *Results of SPiN task performance by Condition. Significant differences were found between Condition 1 and Condition 2 as well as between Condition 1 and Condition 3. No significant difference was found between Conditions 2 and 3. Error bars represent the 95% Confidence Intervals of the performance observations.*

As performance data was binary (correctly or incorrectly recognized) and per trial, a multilevel mixed-effects logistic regression model was used. The mixed-effects logistic regression model took performance as the independent variable and the main effects of block, condition, and their interaction as fixed effects. Additionally, individual subject differences were accounted for by taking participants as a random effect modeled as random intercepts. Initial testing for an interaction between Condition and Block revealed no significant interaction between the two ($p = .0705$), suggesting participant improvement throughout the experiment did not significantly differ between noise pattern conditions, so the interaction was dropped from further analysis.

The final model (with subject as a random intercept and without condition-block interaction) revealed a main effect of condition on performance ($p = .0005$), implying that the type of noise patterns the participant encountered significantly affected performance on a given trial. We also found a main effect of block number on performance ($p < .0001$), implying that there may be learning over the course of the experiment.

For comparison between conditions, the model shows a highly significant difference between Condition 1 and Condition 2 ($\beta = -.213, z = -3.24, p = .001$), pointing to a significant advantage afforded by the repetition of patterns within a trial rather than the repetition of the masking pattern between trials. We also found a highly significant difference between Condition 1 and Condition 3 ($\beta = -.232, z = -3.53, p < .001$), further cementing the indication of the usefulness of repeating patterns within trials. There was no difference found between Condition 2 and Condition 3 however ($\beta = -.019, z = -.29, p = .771$), implying that the repetition of the masking pattern across trials did not improve performance in the task. Now that the findings of the main task have been presented, we will carry out a comparison of participants' performance on the main task to their performance on each of the individual differences tasks.

Participants' performance on the Simon task was measured in the form of delay between conditions of reaction time from the onset of the trial's visual stimulus until a response was provided by mouse click. Participants' mean reaction time delay caused by incongruence was 29ms. The 95% confidence interval for the delays ranged from 21ms to 37ms. Correlations between participants' mean reaction time delay and their performance on the SPiN task was tested using a linear regression model. Participants' overall performances on the SPiN task were not significantly correlated with their mean incongruence delay measured in the Simon task ($\beta = -.097, t(57) = -1.59, p = .118$), indicating that greater executive function ability did not correlate to better performance on the SPiN task. In the SPiN results, the advantage participants were able to glean from within trial noise pattern repetition was measured by taking the difference for each participants' mean performance in Condition 3 (the control condition) and subtracting it from their mean performance in Condition 1. This difference between condition means was found to be uncorrelated to performance on the Simon task ($\beta = -.022, t(57) = -.76, p = .452$), suggesting that improved executive function did not lead to subjects being able to derive more advantage from repeating noise patterns within trials. Now that the task of executive function has been analyzed, the working memory task must be analyzed.

Performance on the Digit Span task was calculated as a simple percentage of total trials correct. Participants' mean performance for ordered digit span recall was 76%. The 95% confidence interval ranged from 72% to 79%. For comparison of performance on the SPiN task (both overall and Condition 1 advantage) to performance on the Digit Span task, a simple linear regression model was used. Overall performance in the SPiN task was not significantly correlated with performance on the Digit Span task ($\beta = 11.2, t(57) = .81, p = .423$), indicating that increased working memory did not aid participants in the SPiN task. The same Condition 1 advantage calculated for each participant above (mean Condition 1 performance minus mean Condition 3 performance) was also tested for correlation to Digit Span task performance. The Condition 1 advantage was not found to be significantly correlated to Digit Span performance ($\beta = 1.8, t(57) = .27, p = .79$), suggesting that increased working memory also did not benefit subjects' ability to utilize the repeating pattern information. Lastly, performance on the Simon task and Digit Span task were checked for correlation using a linear regression model. The model revealed the two to be uncorrelated as well ($\beta = -11, t(58) = -.37, p = .713$), suggesting working memory and executive function were not correlated given our measures. Given these results, we shall now discuss the implications of our findings.

4. Discussion

As mentioned before in the introduction, traditionally, informational masking and energetic masking were consigned to their two separate domains with informational masking operating on the central higher-up processes and energetic masking operating on the peripheral lower-down processes. With the exception of space providing a common release from both types of masking[19][20][21], the interaction between the two noise types and where one's domain ends and another's begins is still largely unknown despite its being quite an old problem[22][23]. While most continue to conceptualize of the two types of masking as separate entities, a few have begun to probe the question of how separate the two really are, and whether they can be combined to create novel masking effects. A variation of the picket fence pattern[24] discussed in the introduction was carried out by [25] in which in addition to interleaving noise and speech, they also interleaved distractor speech and target speech. In this variation, they used both a picket fence pattern, as well as an interrupted target and continuous distractor. They found that participants performed better in the continuous distractor condition than the interleaved condition implying that subjects may be using information from the distractor speech to segregate the target and distractor speech, much as participants in this study may be using the characteristics of the background noise to segregate the masker and target streams. [26] used LTAS and envelope manipulations and found release from a masker with spectral and envelope characteristics of speech compared to a steady-state masker, similar to the release found here using LTAS noise, however the noise used here did not correlate to a speech envelope, but to the glimpsing patterns.

While there has been a few precedents for the manipulation of noise to test energetic masking, the novel paradigm of glimpsing window patterns revealed the possibility that participants were using online temporal information to gain release from energetic masking. In order for this release to occur, several higher order processes are implicated in the process. In order for the online information to be extracted from the patterns and used to provide top-down information, it is plausible that a pattern recognition system is used. In addition to this pattern recognition system, the auditory attention system most likely plays a part while participants are focused on the task. While they are attending to the input in the preamble and the pattern recognition is occurring, it is also possible that some of the mechanisms utilized for auditory stream segregation are constructing a percept of the background noise, and thus any new information in the input, such as the target at the onset of the third pattern, may be in its own perceptual stream, aiding in its perception and comprehension. While these central cognitive processes are often only implicated with informational masking, perhaps the two are not as distinct as once thought.

5. Conclusions

This experiment addressed the question of which mechanisms release from energetic masking could utilize, given temporal regularities in the noise masker. A new paradigm was tested creating patterns of morse code glimpsing windows with jitter, introducing several complex timing structures into the noise signal, something previous studies had not done. Despite these timing complexities, participants performed significantly better on trials in which they had online temporal information available about the patterns than in those where such information was unavailable. Participants are hypothesized to have used the fairly high level processes of pattern recognition, auditory attention, and perhaps auditory streaming, to create top-down information to aid in performance on a speech perception in noise task. This release from energetic masking using a higher level cognitive function also shows that the relegation of energetic masking solely to the periphery warrants reexamination.

6. Acknowledgements

Funding was provided by a USC Graduate School Ph.D. Fellowship and NIH grant T32 DC009975. The author thanks Dr. Jason Zevin and his students at the Cognitive Neuroscience lab.

7. References

[1] Leek, M. R., Brown, M. E., & Dorman, M. F. (1991). Informational masking and auditory attention. Perception & Psychophysics, 50(3), 205214. http://doi.org/10.3758/BF03206743

[2] Scott, S. K., Rosen, S., Wickham, L., & Wise, R. J. S. (2004). A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. The Journal of the Acoustical Society of America, 115(2), 813821. http://doi.org/10.1121/1.1639336

[3] Zhang, C., Lu, L., Wu, X., & Li, L. (2014). Attentional modulation of the early cortical representation of speech signals in informational or energetic masking. Brain and Language, 135, 8595. http://doi.org/10.1016/j.bandl.2014.06.002

[4] Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. The Journal of the Acoustical Society of America, 109(3), 11011109. http://doi.org/10.1121/1.1345696

[5] Durlach, N. I., Mason, C. R., Kidd, G., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. G. (2003). Note on informational masking (L). The Journal of the Acoustical Society of America, 113(6), 2984. http://doi.org/10.1121/1.1570435

[6] Brungart, D. S., Chang, P. S., Simpson, B. D., & Wang, D. (2006). Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation. The Journal of the Acoustical Society of America, 120(6), 4007. http://doi.org/10.1121/1.2363929

[7] Bregman, A. S. (1990). Auditory scene analysis [electronic resource]: the perceptual organization of sound. Cambridge, Mass.: MIT Press.

[8] Bizley, J. K., & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. Nature Reviews. Neuroscience, 14(10), 693707. http://doi.org/http://dx.doi.org.libproxy2.usc.edu/10.1038/nrn3565

[9] Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of Attention on Neuroelectric Correlates of Auditory Stream Segregation. Journal of Cognitive Neuroscience, 18(1), 113. http://doi.org/10.1162/089892906775250021

[10] Shamma, S. A., Elhilali, M., & Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. Trends in Neurosciences, 34(3), 114123. http://doi.org/10.1016/j.tins.2010.11.002

[11] Cusack, R., Decks, J., Aikman, G., & Carlyon, R. P. (2004). Effects of Location, Frequency Region, and Time Course of Selective Attention on Auditory Scene Analysis. Journal of Experimental Psychology: Human Perception and Performance, 30(4), 643656. http://doi.org/http://dx.doi.org.libproxy2.usc.edu/10.1037/0096-1523.30.4.643

[12] Cooke, M. (2003). Glimpsing speech. Journal of Phonetics, 31(34), 579584. http://doi.org/10.1016/S0095-4470(03)00013-5

[13] Cooke, M. (2006). A glimpsing model of speech perception in noise. The Journal of the Acoustical Society of America, 119(3), 1562. http://doi.org/10.1121/1.2166600

[14] Licklider, J. C. R., & Miller, G. A. (1948). The Intelligibility of Interrupted Speech. The Journal of the Acoustical Society of America, 20(4), 593593. http://doi.org/10.1121/1.1916995

[15] Clopper, C. G., & Pisoni, D. B. (2006). The Nationwide Speech Project: A new corpus of American English dialects. Speech Communication, 48, 633-644.

[16] Boersma, Paul & Weenink, David (2012). Praat: doing phonetics by computer [Computer program]. http://www.fon.hum.uva.nl/praat/ (6 December, 2012).

[17] Ng, B. S. W., Schroeder, T., & Kayser, C. (2012). A Precluding But Not Ensuring Role of Entrained Low-Frequency Oscillations for Auditory Perception. The Journal of Neuroscience, 32(35), 1226812276. http://doi.org/10.1523/JNEUROSCI.1877-12.2012

[18] Riecke, L., Sack, A. T., & Schroeder, C. E. (2015). Endogenous Delta/Theta Sound-Brain Phase Entrainment Accelerates the Buildup of Auditory Streaming. Current Biology, 25(24), 31963201. http://doi.org/10.1016/j.cub.2015.10.045

[19] Best, V., Ozmeral, E., Gallun, F. J., Sen, K., & Shinn-Cunningham, B. G. (2005). Spatial unmasking of birdsong in human listeners: Energetic and informational factors. The Journal of the Acoustical Society of America, 118(6), 37663773. http://doi.org/10.1121/1.2130949

[20] Arbogast, T. L., Mason, C. R., & Kidd, G. J. (2002). The effect of spatial separation on informational and energetic masking of speech. The Journal of the Acoustical Society of America, 112(5), 20862098. http://doi.org/10.1121/1.1510141

[21] Ihlefeld, A., & Shinn-Cunningham, B. (2008a). Disentangling the effects of spatial cues on selection and formation of auditory objectsa). The Journal of the Acoustical Society of America, 124(4), 22242235. http://doi.org/10.1121/1.2973185

[22] Miller, G. A. (1947). The masking of speech. Psychological Bulletin, 44(2), 105.

[23] Tanner, W. P. J. (1958). What is Masking? The Journal of the Acoustical Society of America, 30(10), 919921. http://doi.org/10.1121/1.1909406

[24] Miller, G. A., & Licklider, J. C. R. (1950). The Intelligibility of Interrupted Speech. The Journal of the Acoustical Society of America, 22(2), 167173. http://doi.org/10.1121/1.1906584

[25] Iyer, N., Brungart, D. S., & Simpson, B. D. (2007). Effects of periodic masker interruption on the intelligibility of interrupted speech. The Journal of the Acoustical Society of America, 122(3), 16931701. http://doi.org/10.1121/1.2756177

[26] Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. The Journal of the Acoustical Society of America, 88(4), 17251736. http://doi.org/10.1121/1.400247