

Explaining intelligibility in speech-modulated maskers using acoustic glimpse analysis

Bobby E. Gibbs II^{a)} and Daniel Fogerty

*Department of Communication Sciences and Disorders, University of South Carolina,
Columbia, South Carolina 29208, USA
gibbsbe@mailbox.sc.edu, fogerty@sc.edu*

Abstract: Intelligibility was measured in speech-modulated noise varying in level and temporal modulation rate (TMR). Acoustic analysis measured glimpses available above a local signal-to-noise ratio criterion (LC). The proportion and rate of glimpses were correlated with intelligibility, particularly in relation to masker level or TMR manipulations, respectively. Intelligibility correlations for each metric were maximized at different analysis LCs. Regression analysis showed that both metrics measured at -2 dB LC were required to best explain the total variance ($R^2 = 0.49$) for individual sentence intelligibility. Acoustic conditions associated with recognizing speech in complex maskers are best explained using multidimensional glimpse metrics.

© 2018 Acoustical Society of America

[DDO]

Date Received: November 14, 2017 **Date Accepted:** May 22, 2018

1. Introduction

The dips in a modulated masker (MM) provide opportunities for momentary “glimpses” of target speech (e.g., [Cooke, 2006](#)) that are preserved at favorable signal-to-noise ratios (SNRs). The measurement of these glimpses requires selecting a local SNR criterion (LC) above which speech information is sufficient for intelligibility ([Cooke, 2006](#)). These glimpses can then be analyzed using different metrics, such as the proportion of glimpsing (i.e., the ratio of target speech glimpsed compared to total target speech) or rate of glimpsing (i.e., the number of glimpses per second). Several studies using interrupted speech have observed that the proportion of glimpses alone may not be adequate to explain the acoustic conditions contributing to performance in MMs ([Miller and Licklider, 1950](#); [Wang and Humes, 2010](#); [Shafiro *et al.*, 2011](#)). Thus, a multidimensional account of speech glimpses may be necessary to adequately explain the acoustic conditions contributing to speech recognition. Interactions between glimpse properties related to level and modulation rate dimensions are likely common in everyday listening. Currently, the extended speech intelligibility index (ESII; [Rhebergen and Verseld, 2005](#)), successfully accounts for the proportion of glimpses when predicting MM intelligibility. However, the ESII calculates a time-averaged measure of preserved speech across the sentence, which does not consider how the distribution of speech cues may affect intelligibility (e.g., [Buss *et al.*, 2009](#)). The purpose of this study was to evaluate how multiple glimpse metrics may be combined to explain the acoustic conditions contributing to intelligibility in multidimensional speech-based maskers.

Multiple factors related to the distribution of partial speech information were studied by [Wang and Humes \(2010\)](#), who used periodically interrupted speech to investigate intelligibility as a function of the rate of interruption. The proportion of speech preserved was also investigated by varying the on-duration of speech within each interruption cycle (i.e., glimpse duration). Recognition was most determined by the proportion of glimpsed speech. However, the rate and duration of interruptions interacted with the proportion of preserved speech. When speech was highly degraded, recognition improved for shorter, more frequent, glimpses. Conversely, when the proportion of speech preserved was high, recognition was better for longer infrequent glimpses. Similarly, [Shafiro *et al.* \(2011\)](#) reported that the effect of a secondary rate of interruption within a primary rate determined intelligibility differently depending on the overall proportion of glimpsed speech as well as the quality of the cues needed for word identification. These studies of interrupted speech demonstrate a complicated relationship between the proportion of speech preserved and the interruption rate in determining intelligibility.

^{a)} Author to whom correspondence should be addressed.

Recent investigations have examined glimpsing in more natural speech-based maskers (Fogerty *et al.*, 2016; Gibbs and Fogerty, 2016). These studies used a speech-modulated noise time compressed or expanded to produce different modulation rates. Conditions were all presented at a single long-term SNR. Both the proportion of glimpsed speech and the rate of glimpses were correlated with intelligibility. However, the association of either metric with intelligibility was highly dependent on the chosen LC (Gibbs and Fogerty, 2016), that is, acoustic glimpse properties used during speech recognition may be best captured at different LCs.

The current study evaluated how different glimpse metrics, in isolation or combined, define the acoustic conditions that best explain speech recognition for maskers that vary in level and modulation rate. Furthermore, these metrics were measured across a range of LCs to understand the capabilities and limitations listeners have in glimpsing speech based on these properties.

2. Perceptual analysis of glimpsing in speech-modulated noise

Speech intelligibility was measured in speech-modulated noise at three different global SNRs, i.e., the average SNR across the sentence. The temporal modulation rate (TMR) of the noise was varied using time expansion/compression. Subsequent glimpse analyses were then used to explain performance.

2.1 Methods

Five listeners participated in the experiment (mean age = 23.4 yr, standard deviation = 1.9 yr). All participants had audiometric thresholds less than 20 dB hearing level at octave frequencies from 250 to 8000 Hz. While listeners had participated in similar experiments, none were familiar with the sentences tested.

Stimuli consisted of IEEE sentences spoken by a male talker. To generate the speech-modulated masking noise, 40 IEEE sentences (not used in the final stimulus presentation) were first concatenated into a single file and the silences were removed between the sentences. Next, a steady-state speech-shaped noise (SSN) was created to match the long-term average spectrum of the speech concatenation. The speech concatenation duration was time expanded/compressed using Pitch Synchronous Overlap and Add to 25%, 100%, or 400% of the original duration. This manipulation effectively multiplied the masker TMR by a factor of 4 (fast TMR), 1 (natural TMR), or 0.25 (slow TMR), respectively. Through half-wave rectification, the temporal envelope of each modified speech concatenation was extracted and then low-pass filtered using a sixth-order Butterworth filter to preserve modulations up to 16 Hz. Each extracted temporal envelope was then used to amplitude modulate the SSN, creating three speech-modulated noise maskers that varied according to three different TMRs. For presentation, a random segment of one of the maskers was added to the target (matching its overall duration) to create the final stimulus. Both target and masker speech sentences were low-pass filtered to 6400 Hz to match the bandwidth used in previous studies (Fogerty *et al.*, 2016; Gibbs and Fogerty, 2016). Speech and masker files were saved in separate channels for later glimpse analysis. In the final presentation, speech and noise channels were played concurrently with the noise turning on and off with the speech.

To assess the effect of global SNR on glimpsing, three global SNRs (−8, −4, and 0 dB SNR) were chosen based on the long-term root-mean-squared (RMS) of the target sentence relative to the noise. There were 30 sentences in each block for a total of 270 sentences presented with modulated-noise (30 sentences × 3 TMRs × 3 SNRs). An additional three blocks of 10 sentences each, one for each of the SNR conditions, were tested in the presence of the SSN for a combined total of 300 sentences.

The listening tests took place in a sound-attenuating booth. Speech was calibrated to 70 dB sound pressure level and presented through the right channel of Sennheiser (Weddenmark, Germany) HD 280 Pro headphones. The experiment was conducted using a MATLAB interface for self-paced presentation. Within each global SNR block, stimuli with different TMRs were randomly interspersed. The order of the SNR presentation blocks was randomly chosen for each participant. Participants were instructed to repeat aloud each sentence and encouraged to guess. Responses were recorded and subsequently scored offline. Participants completed a practice session prior to the experimental task that consisted of nine sentences (two from each of the three TMRs and three additional sentences presented in SSN) at an intermediate global SNR of −4 dB.

2.2 Results and discussion

A sentence-level analysis was conducted so that accuracy and glimpse metrics could be explicitly compared for each sentence. Keyword accuracy was converted to rationalized

arcsine units (RAUs) and averaged across listeners for each sentence. The average accuracy across sentences for each experimental condition is displayed in Fig. 1(A). Participants typically gained over 40 RAU points in accuracy for MM compared to SSN when tested at negative SNRs. A poorer performance for MM compared to SSN at 0 dB global SNR occurred for the slowest TMR, which suggests that very slow amplitude modulations may interfere with intelligibility even at a favorable global SNR.

To explore intelligibility in MM conditions, a 3 (global SNR: -8, -4, 0 dB) × 3 (TMR: fast, natural, slow) analysis of variance was conducted on the listener-averaged RAU scores obtained for each sentence ($N=270$). Overall, significant main effects were observed for global SNR [$F(2,261)=105.19, p<0.001, \eta_p^2=0.45$] and TMR [$F(2,261)=41.15, p<0.001, \eta_p^2=0.24$], as well as a small interaction effect [$F(4,261)=2.57, p=0.038, \eta_p^2=0.04$].

Significant main effects were investigated more thoroughly through planned t -test comparisons using the Bonferroni adjustment. For global SNRs of -4 and -8 dB, the fast TMR was significantly ($p<0.001$) higher than other rates by 25–46 RAU points. Accuracy was not significantly different between natural and slow TMRs for global SNRs of -8 dB ($p=1.0$) or -4 dB ($p=0.65$). At a global SNR of 0 dB, significant differences existed between slow and fast TMRs ($p<0.001$) and between slow and natural TMRs ($p=0.006$), but not between fast and natural TMRs ($p=0.58$).

Overall, both global SNR and TMR exerted significant effects on intelligibility. Fast TMR resulted in a better performance at poorer global SNRs. Slow TMR resulted in a poorer performance at the most favorable global SNR (0 dB) when long “on” periods of the MM still resulted in negative local SNRs.

3. Acoustic analysis of glimpses

The first analysis investigated how individual glimpse metrics were able to explain variability in performance within each fixed SNR-TMR condition. The second analysis examined correlations across all stimulus conditions to examine how either of these glimpse metrics explained performance as noise backgrounds varied both in SNR and TMR dimensions. In the final analysis, multiple linear regression was used to examine how glimpse metrics can be combined to better capture performance in these multidimensional noise conditions.

3.1 Glimpse analysis methods

For each target-masker signal, the running RMS (dB) was obtained to compute the short-time SNR using 16-ms non-overlapping windows. Glimpses were therefore defined as wideband temporal intervals of speech occurring above the LC for at least 16 ms. This simplistic approach isolates the effect of temporal glimpsing. A frequency-based approach would likely yield a more nuanced account of glimpses that listeners have access to. All measures were calculated across the entire stimulus file. Two metrics for characterizing glimpses were used:

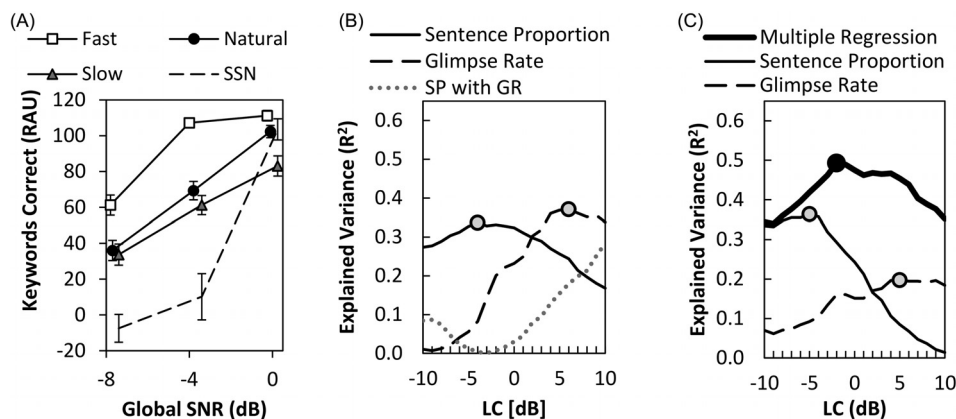


Fig. 1. (A) Average keyword accuracy (in RAU) across sentences. Separate lines indicate separate TMR conditions. The dashed line indicates performance in SSN. Error bars represent one standard error of the mean. (B) Squared bivariate correlations between intelligibility (RAU) and either SP (solid line) or GR (dashed line). Correlations are plotted across local criteria (LC) used in the analysis. The circled points indicate the maximum correlation obtained for either metric. The gray dotted line shows the correlation between the SP and GR metrics. (C) The bold black line shows R^2 output for multiple regression using both metrics measured at the same LC to predict accuracy. The thin black lines display the squared semi-partial correlations between intelligibility and either SP ($R^2_{RAU(SP,GR)}$, thin solid line) and GR ($R^2_{RAU(GR,SP)}$, thin dashed line) for the multiple regression analysis.

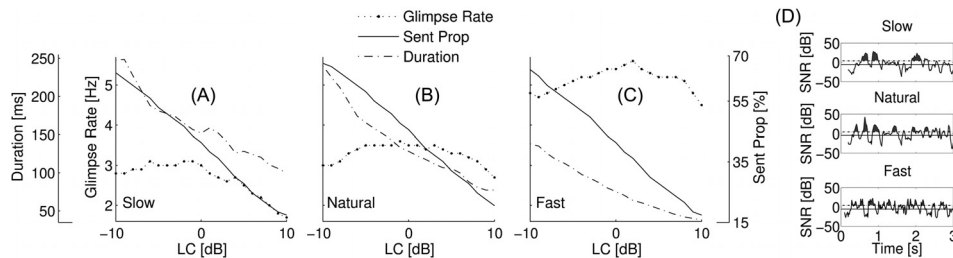


Fig. 2. Glimpse metric relationships for sentences in the 0 dB global SNR condition as a function of LC. Glimpse metrics are averaged for sentences in each TMR condition. Panels (A)–(C): Glimpse metrics are displayed due to varying LC while holding TMR constant (equivalent to changing the global SNR) or with LC held constant across the different TMR panels. Panel (D) displays short-term SNR plots for an example sentence with random masker noise added based on the TMR condition. Glimpsed time intervals are shaded based on an LC of 4 dB (dashed line) to show available glimpses at an intermediate global SNR of -4 dB. Alternative glimpse intervals provided by an LC of -4 dB can be observed below the dashed line.

- (1) *Sentence Proportion (SP)*: The proportion of the stimulus file that occurred above the LC.
- (2) *Glimpse Rate (GR)*: The average number of glimpses per second—derived by dividing the total number of glimpses across the stimulus by the stimulus file duration.

Figure 2 displays the relationship between glimpse measures for the sentences and noise conditions tested. SP, GR, and glimpse duration (i.e., the average duration of individual glimpses) were calculated for sentences measured at a global SNR of 0 dB as LC varied for the three different TMRs. Glimpse metrics were averaged across the sentences for each TMR and plotted across LCs from -10 to 10 dB. For a given TMR [i.e., panel (A), (B), or (C)], as the LC (or SNR) changes, so do SP and glimpse duration, while GR remains relatively constant (except at extreme positive LCs which decrease the number of glimpses observed or at extreme negative LCs which merge adjacent glimpses together). This relationship can be observed in panel (D) which displays the running SNR for each of the three TMRs. As LC increases (i.e., from the solid to the dotted line) the glimpses (i.e., shaded regions) are similarly frequent but they become shorter (i.e., reduced glimpse duration) and the overall proportion of glimpsed speech across the entire waveform is also reduced. In contrast, for a given LC (or SNR), as the TMR is increased [i.e., compare across panels (A)–(C)], SP remains relatively constant while GR increases and glimpse duration decreases. Inspection of the waveforms in panel (D) also demonstrates this relationship. As the TMR is increased [i.e., going from top to bottom in panel (D)], the glimpses become more frequent and shorter in duration, yet the overall proportion glimpsed remains constant.

This analysis demonstrates that glimpse duration co-varies with GR and SP. As a result, it is not independently predictive of intelligibility in speech MMs (Gibbs and Fogerty, 2016). Therefore, the current analysis focused on the contribution of SP and GR, which other studies have reported to exhibit unique contributions to intelligibility (e.g., Wang and Humes, 2010).

3.2 Results

Glimpse analysis for each combination of masker level and TMR. Correlations between a given glimpse metric and RAU sentence accuracy (averaged across listeners) were analyzed for each of the nine combinations of fixed global SNR and TMR (see Table 1). This analysis isolates the effect of glimpse properties on performance in different conditions while holding masker properties constant. The optimal LC for each

Table 1. Maximum glimpse metric correlations with intelligibility are displayed for each global SNR by TMR combination. LCs (in dB) associated with maximum correlations are indicated in parentheses. Bold indicates $p < 0.05$.

dB SNR	Fast		Natural		Slow	
	SP	GR	SP	GR	SP	GR
-8	0.24 (7)	0.29 (4)	0.56 (-2)	0.60 (3)	0.57 (-1)	0.57 (5)
-4	-0.29 (7)	0.46 (0)	0.36 (5)	0.35 (-5)	0.45 (-1)	0.71 (4)
0	0.20 (-2)	0.36 (-2)	0.63 (5)	-0.59 (-6)	0.29 (-10)	0.46 (-5)

metric was identified based on the LC that resulted in the maximal intelligibility correlation for that condition across an LC range of -10 to 10 dB. Of the nine conditions, only one (fast TMR at -8 dB SNR) was not significantly associated with either SP or GR. Inspection of the other eight conditions suggests that GR was most effective at capturing variation within the fast TMR conditions, while a combination of SP and GR were effective at capturing performance for the natural and slow TMR conditions. This analysis indicates that both glimpse metrics can be effective at capturing variation in performance associated with changing glimpse properties. However, variable associations are observed depending on the specific SNR-TMR combination.

Glimpse analysis across all masker conditions. This next glimpse analysis was conducted to determine how individual glimpse metrics explained performance in multidimensional acoustic backgrounds varying both in masker level and TMR. The correlation of each metric with intelligibility across all MM conditions is displayed in Fig. 1(B) for an LC range of -10 to 10 dB. Due to the pooling across all masker conditions, this analysis examines glimpse metric correlations that could be driven by both masker properties: global SNR and TMR. The gray dotted line displays the correlation between the SP and GR metrics. While GR and SP evidenced similar maximum intelligibility correlations, the LCs associated with these maximums are distinct (GR: 6 dB LC, $R^2=0.37$; SP: -4 dB LC, $R^2=0.34$). The two metrics show equivalent correlations at an LC of 2 dB. Below this point SP shows stronger correlations than GR, with the reverse true above 2 dB LC.

Consideration of these functions helps to explain how listeners may use speech cues distributed across the sentence. The peak in the SP correlation function occurred at -4 dB LC [Fig. 1(B), solid line]. This negative LC suggests that listeners derive a benefit from somewhat noisy glimpses in agreement with Cooke (2006). Furthermore, increases in the total duration of these noisy glimpses, which occurs due to global SNR improvement, results in improvements in intelligibility. In contrast, the GR correlation function peaked at 6 dB LC [Fig. 2(A), dashed line]. This finding suggests that availability of relatively short, pristine glimpses that regularly sample the information across the sentence (i.e., “multiple looks”) are beneficial. According to Buss *et al.* (2009), this result might not occur for more predictable speech materials that require less frequent sampling across the sentence. The present results suggest that some speech cues may be robust in noise while others may be more susceptible to masking, suggesting that analysis may require multiple LCs.

To summarize, across all MM conditions both glimpse metrics were correlated with intelligibility. However, these metrics appear to index different effects of the MM on intelligibility. These results are consistent with Wang and Humes (2010) who suggested that, among other factors, the contribution of multiple looks combined with the total proportion of the speech signal is crucial for intelligibility.

Complicating this interpretation of multiple factors is that glimpse metrics may be correlated, especially at positive LCs. Therefore, it is not clear whether variance in the data is adequately captured by either metric alone or if better performance is obtained by considering both metrics together.

Combined glimpse metric analysis. The purpose of this final glimpse analysis was to quantify how combining multiple glimpse metrics might aid in explaining speech intelligibility in multidimensional speech-modulated backgrounds. Multiple linear regression was used to investigate the combined contribution of GR and SP in predicting average listener intelligibility for individual sentences. These regression models were further examined to clarify the independent contribution of each glimpse metric in explaining variance in the data.

An initial linear regression model predicted RAU accuracy across all MM conditions from GR and SP metrics measured at maximal LCs identified previously (6 dB for GR; -4 dB for SP). Thus, different LCs were used to measure the two glimpse metrics, based on the squared bivariate correlation functions [Fig. 1(B)]. This model explained 47.6% of the variance [$F(2,267) = 121.13$, $p < 0.001$]. Multicollinearity, assessed by a variance inflation factor (VIF) was low for predictors in this model (VIF = 1.31).

Next, we investigated the variance explained in MM using glimpse metrics measured at the same LC. For this analysis, a range of LCs (-10 to 10 dB) was tested, with separate models for each LC, to determine the optimal LC for the combined function. The R^2 values for these regression models are plotted in the bold solid line of Fig. 1(C). The best-fitting model occurred at an LC of -2 dB. Here 49.3% of the variance was explained [$F(2,267) = 129.62$, $p < 0.001$]. This is likely due to reducing the correlation between glimpse metrics, accounting for more unique variance.

Multicollinearity was indeed reduced from the initial model ($VIF = 1.01$). This more restricted model, using only a single LC value, explained more of the variance.¹

To explore the independent contributions of each glimpse metric within the regression models, squared semi-partial correlations were computed [thin lines in Fig. 1(C)]. The squared semi-partial correlation can be interpreted as the proportion of unique variance accounted for by the predictor variable relative to the total variance in the dependent variable. Note that at negative LCs, the unique variance accounted for by SP remains high. The function maximum for the semi-partial correlation is close to that of the bivariate correlation for SP (i.e., -5 vs -4 dB LC, respectively). At positive LCs, GR still explains the majority of the variance in RAU accuracy, but the unique variance explained by this metric is reduced after accounting for the shared variance with SP. Again, the semi-partial function maximum is close to the bivariate function for GR (i.e., 5 vs 6 dB LC, respectively). For the best-fitting model at -2 dB LC the squared semi-partial correlation for SP was 0.29 and for GR was 0.16. Contrast these results with the initial regression model (based on different LCs) where squared semi-partial correlations were 0.10 and 0.14 for SP and GR, respectively. Clearly a more unique variance was accounted for by the predictors in the best-fitting model in which a single LC was used in the analysis. For this model, SP accounted for more of the unique variance in accuracy than the GR, but neither glimpse metric alone was able to adequately explain performance. Therefore, this analysis confirms the results of the independent analysis: a combination of glimpse factors are critical for explaining listener performance in MM.

A limitation of this analysis is the observation from Fig. 2 that SP is associated with changes in the global SNR, while the GR is associated with changes in TMR. This association is necessary for these acoustic measures of the speech-masker mixture to capture variation in performance associated with the different conditions. Indeed, one reason the ESII is successful is that it uses changes in the SNR to predict performance. However, as these glimpse metrics were able to explain some variance in performance within each global SNR-TMR combination (see Table 1), a question of interest is the degree to which these measures are able to explain performance beyond what can be explained by knowing the SNR and TMR parameters of the experimental conditions. To address this, the residuals were calculated from multiple linear regression with global SNR and TMR as predictor variables for the listener-averaged RAU scores. Global SNR explained 36% of the variance with TMR explaining an additional 11% for a combined total of 47% of the variance [$F(2,267) = 117.7$, $p < 0.001$]. GR and SP at an LC of -2 dB were then entered into stepwise linear regression as predictors for this residual score (after global SNR and TMR were accounted). In this final analysis, only the GR was a significant predictor accounting for an additional 6% of the variance beyond that accounted for by global SNR and TMR parameters [$F(1,268) = 15.5$, $p < 0.001$].² These results suggest that global SNR combined with TMR provides a reasonable account of the acoustic conditions that contribute to average listener performance. The perceptual consequences of these experimental parameters are accurately captured by measuring GR and SP. Furthermore, additional variance is explained by the GR alone, suggesting that acoustic analysis of the distribution of glimpses should be considered to fully account for the acoustic conditions governing sentence intelligibility in MM.

4. General discussion

While this study was limited to only wideband temporal glimpse analysis, the results suggest that the selection of a glimpse metric and corresponding LC significantly affect the ability to associate acoustic masker properties with intelligibility. By remaining sensitive to glimpse phenomena, a model such as ESII makes reasonably accurate predictions of recognition in MMs (Rhebergen and Verseld, 2005), even when considering complex maskers (Rhebergen *et al.*, 2008). However, the ESII is not designed to differentiate between benefits in glimpsing due to the distribution of glimpses. The current analysis found that the greatest amount of variance in accuracy across speech MMs was best explained by a combination of increases in the availability of longer somewhat noisy glimpses and more frequent sampling of shorter glimpses across the sentence (indexed by the GR and SP measured at -2 dB LC). These results suggest that a combination of factors is required to describe how listeners glimpse distributed speech cues for sentences in speech-modulated noise. Further research is required to examine if the ESII might improve intelligibility predictions by also considering properties of the GR.

Acknowledgments

This work was supported, in part, by grants from the National Institutes of Health, National Institute Deafness and Other Communication Disorders, Grant Nos. R03-DC012506 and R01-DC01565.

References and links

¹A regression was run with intelligibility averaged across listeners for each IEEE list and associated glimpse metrics measured across lists for each condition to control for linguistic variability and random noise across sentences. The pattern of variance explained as a function of LC was the same as the sentence-level analysis with an LC of -2 dB explaining 89% of the variance in the data [$F(2,24) = 100.25, p < 0.001$].

²A similar analysis was also run for list scores (instead of sentence scores). Multiple linear regression with global SNR and TMR combined accounted for 82.8% of the variance for lists scores [$F(2,24) = 57.6, p < 0.001$]. The GR accounted for an additional 16.5% of the variance on the remaining residual scores [$F(1,25) = 5.0, p < 0.05$] with no additional significant variance accounted for by SP.

Buss, E., Whittle, L. N., Grose, J. H., and Hall, J. W., III. (2009). "Masking release for words in amplitude-modulated noise as a function of modulation rate and task," *J. Acoust. Soc. Am.* **126**, 269–280.

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

Fogerty, D., Xu, J., and Gibbs, B. E., II. (2016). "Modulation masking and glimpsing of natural and vocoded speech during single-talker modulated noise: Effect of the modulation spectrum," *J. Acoust. Soc. Am.* **140**, 1800–1816.

Gibbs, B. E., II, and Fogerty, D. (2016). "Glimpsing predictions for natural and vocoded sentence intelligibility during modulation masking: Effect of the glimpse cutoff criterion," in *INTERSPEECH 2016*, pp. 1677–1681.

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

Rhebergen, K. S., and Versfeld, N. J. (2005). "A speech intelligibility index-based approach to predict the speech reception threshold for sentences in fluctuating noise," *J. Acoust. Soc. Am.* **117**, 2181–2192.

Rhebergen, K. S., Versfeld, N. J., and Dreschler, W. A. (2008). "Prediction of the intelligibility for speech in real-life background noises for subjects with normal hearing," *Ear Hear.* **29**, 169–175.

Shafiro, V., Sheft, S., and Risley, R. (2011). "Perception of interrupted speech: Effects of dual rate gating on the intelligibility of words and sentences," *J. Acoust. Soc. Am.* **130**, 2076–2087.

Wang, X., and Humes, L. (2010). "Factors influencing recognition of interrupted speech," *J. Acoust. Soc. Am.* **128**, 2100–2111.