

# 1

## Experiment 2: Semantic predictability facilitates comprehension of degraded speech in a graded manner

### 1.1 Background

In the previous chapter, we showed that semantic predictability facilitates comprehension of degraded speech. There was a significant difference in word recognition accuracies between low and high predictability sentences (i.e., contextual facilitation), when listeners were instructed to attend to the entire sentence including the context. In this chapter, we examine if semantic predictability facilitates comprehension of degraded speech in a graded manner; whether prediction is probabilistic, or if it is an all-or-none phenomenon. To do so, we determine if listeners' are able to decode the context, at all instances, when they are attending to the entire sentence. We also examine if such facilitatory effect is influenced or modulated by adaptation to degraded speech.

#### 1.1.1 Nature of predictability: Probabilistic or all-or-none?

There are two thoughts, or debate about the nature of predictability effects. One line of studies argue that prediction is all-or-none and deterministic [... CITE...]. For

example, in garden path phenomenon, a parser first tries to predict the simplistic structure of a sentence. If this parsing is disconfirmed by the bottom-up input then the parser backs off and restarts – it reinterprets the context and predicts a new sentence structure. Therefore, according to XYZ, how a parser resolves garden path sentences is an evidence for all-or-none prediction.

Humans use only one of many possible continuations in a sentence. XYZ argue that it is metabolically expensive to predict many parallel linguistic units when only most of them are going to be discarded. Therefore, evolution would not pick a language processing system that predicts multiple parallel linguistic units. Another line of studies show that predictions are probabilistic, or graded. Nieuwland et al. [1] showed in a large-scale replication study of DeLong, Urbach, and Kutas [2] that the N400 amplitude at the sentence-final noun is directly proportional to its cloze probability across a range of high- and low-cloze words. Heilbron et al. [3] also showed that a probabilistic prediction model outperforms a constrained guessing model, suggesting that linguistic prediction is not limited to highly predictable sentence endings, but it operates broadly in a wide range of probable sentence endings. The studies mentioned above were either reading studies, or were conducted with clean speech.

To our knowledge, only one study by Strauß, Kotz, and Obleser [4] has directly studied the nature of prediction in degraded speech comprehension. Severely degraded (4 channels noise vocoded), moderately degraded (8 channels noise vocoded), and clear (non-degraded) speech were presented to the participants. Target word predictability was varied by manipulating its expectancy (i.e., how expected the target word is given the verb) and typicality (i.e., co-occurrence of target word and the preceding verb). Strauß, Kotz, and Obleser [4] reported that at a moderate level of spectral degradation, N400 responses at strong-context, low-typical words and weak-context, low-typical words were largest. N400 responses at the latter two were not statistically different from each other. However, the N400 response was smallest at highly predictable (strong-context, high-typical) words. The authors interpreted these findings as a facilitatory effect of sentence

predictability which might be limited to only highly predictable sentence endings at a moderate level of spectral degradation. Based on these findings, Strauß, Kotz, and Obleser [4] proposed an ‘expectancy searchlight model’. According to the expectancy searchlight model, listeners form ‘narrowed expectations’ from a restricted semantic space when the sentence endings are highly predictable. For less predictable sentence endings, listeners cannot preactivate those less predictable sentence endings in an adverse listening condition. Therefore, the expectancy searchlight model of Strauß, Kotz, and Obleser [4] is in line with all-or-none prediction account. These theoretical accounts have not been further explored in degraded speech comprehension. However, it has been shown that listeners rely on context when speech is moderately degraded [CITE].

### **1.1.2 Predictability effects in degraded speech comprehension**

When the bottom-up perceptual input is difficult to understand, listeners rely more on top-down predictions in adverse listening conditions like noisy environment with background noise [... CITE...], reverberation [... CITE...], and degraded speech [e.g., 5, 6].

In their studies, Obleser and colleagues [7–9], used sentences of two levels of semantic predictability (high and low) and systematically degraded the speech signal by passing it through various numbers of noise vocoding channels ranging from 1 to 32 in a series of behavioral and neuroimaging studies. They found that semantic predictability facilitated language comprehension at a moderate level of speech degradation (at 4 channels, or 8 channels noise vocoding). That is, participants relied on the sentence context when the speech signal was degraded but intelligible enough. Accuracy of word recognition was found to be higher for highly predictable target words than for lowly predictable target words at such moderate levels of speech degradation [8]. For the extremes, i.e., when the speech signal was highly degraded or when it was clearly intelligible, the word recognition accuracy was similar across both levels of sentence predictability, meaning that predictability did

not facilitate language comprehension. It can be safely concluded from these findings that at moderate levels of degradation, participants rely more on the top-down prediction generated by the sentence context and less on the bottom-up processing of unclear, less intelligible (but intelligible enough), and degraded speech signal [10]. In other words, reliance on prediction results in higher word recognition accuracy for high-cloze probability target words than for low-cloze probability target words. In the case of a heavily degraded speech signal, participants may not be able to understand the sentence context and, therefore, they are unable to predict the target word; or their cognitive resources may already be occupied by decoding the signal, leaving little room for making predictions [... CITE... look Ryskin2021]. Thus, there is no differential effect of levels of sentence predictability. On the other extreme, when the speech is clear and intelligible (at the behavioral level, i.e., when the participants respond what the target word of the sentence is), participants recognize the intelligible target word across all levels of sentence predictability. Hence, no differential effect of levels of predictability of target word can be expected.

In contrast to clear speech perception, listeners adapt to degraded speech and their performance has been shown to improve over the course of experiment [e.g., CITE]. Therefore, *adaptation* has to be considered when we review and examine predictability effects on degraded speech.

### 1.1.3 Adaptation to degraded speech

Listeners quickly adapt to novel speech with artificial acoustic distortions [e.g., 11]. Repeated exposure to degraded speech leads to improved comprehension over time [for a review, 12, 13]. When the noise condition is constant throughout the experiment, listeners adapt to it and the performance (e.g., word recognition) improves with as little as 20 minutes of exposure [e.g., 14]. For example, Davis et al. [15, Experiment 1] presented listeners with only 6 channels noise vocoded sentences and found an increase in the proportion of correctly reported words over the course of experiment. Similarly, Erb et al. [16] presented participants with only 4 channels noise vocoded sentences and reached a similar conclusion. In these

experiments, a single noise condition (6 channels or 4 channels) was presented in one block. Therefore, from the participants’ perspective, the next-trial speech degradation level was predictable. Additionally, target word predictability of the sentences were not varied by any measures.

When multiple noise conditions are presented in a (pseudo-)randomized order within a block then a listener is uncertain about any upcoming trial’s noise condition, i.e., if such multiple levels of degradation are due to presentation of multiple channels of noise vocoded speech, then the global channel context is unpredictable or uncertain. This can potentially influence perceptual adaptation. For instance, Mattys et al. [17] note the possibility of total absence of perceptual adaptation, when the characteristics of auditory signal change throughout an experiment. We also know from Sommers, Nygaard, and Pisoni [18] that trial-by-trial variability in the characteristics of distorted speech impairs word recognition [see also, 19]. We thus speculated that if the noise vocoded speech varies from one trial to the next then the adaptation to noise in this scenario might be different from the earlier case in which the noise condition is constant throughout the experiment. Perceptual adaptation, however, is not limited to trial-by-trial variability of stimulus property. Listeners can adapt to auditory signal at different time courses and time scales [20, see also, 21]. In addition to differences in intrinsic trial-by-trial variability and resulting short timescale trial-by-trial adaptation in two channel contexts, the global differences in the presentation of vocoded speech can result in a difference in the general adaptation at a longer timescale between predictable and unpredictable channel contexts.

There is a limited number of studies that has looked at how next-trial noise-uncertainty and global context of speech property influence adaptation. For example, words were presented at +3 dB SNR and +10 dB SNR in a word recognition task in a pseudorandom order [22]. The authors wanted to minimize the certainty about the noise conditions in the block. The same group of authors [23–25] proposed that an adaptive control system (cingulo opercular circuit) might be involved to optimize task performance when listeners are uncertain about upcoming trial. However, we cannot make a firm conclusion about perceptual adaptation *per se*

from their studies as they do not report the change in performance over the course of experiment. Similarly, Obleser and colleagues [7, 9, 26] also presented listeners with noise vocoded sentences (ranging from 2 to 32 channels of noise vocoding) in a pseudo-randomized order but did not report presence or absence of perceptual adaptation to noise vocoded speech. In the abovementioned studies, the authors did not compare participants’ task performance in the blocked design against the presentation in a pseudorandomized block of different noise conditions to make an inference about general adaptation to degraded speech at a longer timescale. To examine the influence of uncertainty about next trial speech features and the global context of speech features on perceptual adaptation, we therefore compared language comprehension with a trial-by-trial variation of sentence predictability and speech degradation either in blocks, in which the noise vocoded channels were blocked, or in a randomized order.

## 1.2 Materials and Methods

### 1.2.1 Participants

We recruited two groups of participants via Prolific Academic and assigned them to one of the two groups: *unpredictable channel context* (n=48;  $\bar{x} \pm \text{SD} = 24.44 \pm 3.5$  years; age range = 18-31 years; 16 females) and *predictable channel context* (n=50;  $\pm \text{SD} = 23.6 \pm 3.2$  years; age range = 18-30 years; 14 females). All participants were native speakers of German residing in Germany. Exclusion criteria for participating in this study were self-reported hearing disorder, speech-language disorder, or any neurological disorder. All participants received monetary compensation for their participation. The study was approved by Deutsche Gesellschaft für Sprachwissenschaft (DGfS) Ethics Committee, and the participants provided consent in accordance with the Declaration of Helsinki.

### 1.2.2 Stimuli

We used the same stimuli described in Section X.X.X in Chapter X.X. The stimuli were digital recordings of 360 German sentences spoken by a female native speaker of German in a normal rate of speech. All sentences were in present tense consisting of pronoun, verb, determiner, and object (noun) in the Subject-Verb-Object form. We used 120 nouns to create three categorizes of sentences – high predictability sentences (HP sentences), medium predictability sentences (MP sentences) and low predictability sentences (LP sentences) – that differed in cloze probability of sentence final target words. (See Appendix A for examples.) Their mean cloze probabilities were  $0.022 \pm 0.027$  ( $\bar{x} \pm \text{SD}$ ; range = 0.00 - 0.09) for LP sentences,  $0.274 \pm 0.134$  ( $\bar{x} \pm \text{SD}$ ; range = 0.1 - 0.55) for MP sentences, and  $0.752 \pm 0.123$  ( $\bar{x} \pm \text{SD}$ ; range = 0.56 - 1.00) for HP sentences. The distribution of cloze probability across LP, MP and HP sentences are shown in Figure X.X.

In the unpredictable channel context, each participant was presented with 120 unique sentences: 40 HP, 40 MP and 40 LP sentences. Channel condition was also balanced across each sentence type, i.e., in each of HP, MP, and LP sentences, ten sentences passed through each noise vocoding channels – 1, 4, 6, and 8 – were presented. This resulted into 12 experimental lists. The sentences in each list were pseudo-randomized, that is, not more than 3 sentences of same noise condition (i.e., same noise vocoding channel), or same predictability condition appeared consecutively. This randomization confirmed the uncertainty of next-trial speech quality (or degradation) in the global context of the experiment.

The same set of stimuli and experimental lists were used in the predictable channel context. Each participant was presented with 120 unique sentences blocked by channel conditions, i.e., blocked by noise vocoding channels. There were four blocks of stimuli. Thirty sentences were presented in each of the four blocks. In the first block, all sentences were 8 channels noise vocoded, followed by blocks of 6 channels, 4 channels, and 1 channel noise vocoded speech consecutively [5]. Within each block, 10 HP, 10 MP and 10 LP sentences were presented. All the

sentences were pseudo-randomized so that not more than three sentences of the same predictability condition appeared consecutively in each block. This ascertained there was a certainty of next-trial speech quality (within each block) and an uncertainty of next-trial sentence predictability across all four blocks.

### 1.2.3 Procedure

Participants were asked to use headphones or earphones. A prompt to adjust loudness was displayed at the beginning of the experiment: A noise vocoded sound not used in the main experiment was presented, and participants were asked to adjust the loudness at their level of comfort. One spoken sentence was presented in each trial. Eight practice trials were presented before presenting 120 experimental trials. They were asked to enter what they had heard (i.e., to type in the entire sentence) *via* keyboard. Guessing was encouraged. The response was not timed. The experiment was about 40 minutes long.

## 1.3 Analyses

In the sentences used in our experiment, verbs evoke predictability of the sentence-final noun. Therefore, the effect of predictability (evoked by the verb) on language comprehension can be rightfully measured if we consider only those trials in which participants identify the verbs correctly. Verb-correct trials were considered as the sentence in which participants realized the context independent of whether they correctly understood the sentence final target noun. Morphological inflections and typos were considered as correct. We first filtered out those trials in which verbs were not identified correctly, i.e., trials with incorrect verbs. Therefore, we included 2469 out of 5760 trials in unpredictable channel context and 2374 out of 6000 trials in predictable channel context from the analyses. The 1 channel noise vocoding condition was dropped from the analyses as there were no correct responses in any of the trials in this condition.



We preprocessed and analysed data in R-Studio (Version 3.6.1; R Core Team, 2019) following the procedure described in Chapter 4.4.4. Response accuracy was analyzed with Generalized Linear Mixed Models (GLMMs) with lme4 [27] and lmerTest [28] packages. Binary responses (correct/incorrect) for all participants in both groups (predictable channel context and unpredictable channel context) were fit with a binomial logistic mixed-effects model. Noise condition (categorical; 4, 6, and 8 channels noise vocoding), target word predictability (categorical; HP, MP, LP), global channel context (categorical; predictable channel context and unpredictable channel context), and the interaction of noise condition and target word predictability were included in the fixed effects.

We first fitted a model with maximal random effects structure that included random intercepts for each participant and item [29]. Both, by-participant and by-item random slopes were included for noise condition, target word predictability and their interaction. To find the optimal model for the data, non-significant higher-order interactions were excluded from the fixed-effects structure (and from the random-effects structure) in a stepwise manner. Model selection was based on Akaike Information Criterion (AIC) [30, 31] unless otherwise stated. Random effects not supported by the data that explained zero variance according to singular value decomposition were excluded to prevent overparameterization. This gave a more parsimonious model [32]. Such a model was then extended separately with: i) item-related correlation parameters, ii) participant-related correlation parameter, and iii) both item- and participant-related correlation parameters. The best fitting model among the parsimonious and extended models, based on AIC, was then selected as the optimal model for our data.

We applied treatment contrast for noise condition (8 channels as a baseline; factor levels: 8 channels, 4 channels, 6 channels) and sliding difference contrast for target word predictability (factor levels: MP, LP, HP) and channel context (factor levels: unpredictable, predictable). The results from the optimal model are shown in Table X.X.X, and are reported below in the Results section.

## 1.4 Results

In this experiment, we tested i) whether predictability facilitates language comprehension only at a moderate level of spectral degradation, and ii) whether adaptation to degraded speech influences language comprehension. We observed that the mean response accuracy increased with an increase in number of noise vocoding channels from 4 to 6 to 8, and with an increase in target word predictability from low to medium to high (see Figure X.X). This trend is consistent across both the channel contexts; Figure X.X and Figure Y.Y show this trend for predictable channel context (i.e., blocked design) and unpredictable channel context (i.e., randomized design) respectively. Mean accuracies across all conditions are given in Table X and Y, and Figure X.X.

These observations are confirmed by the results of statistical analyses. We found a significant main effect of channel condition indicating that the response accuracy was higher in the 8 channels than in the 4 channels ( $\beta = -2.87$ ,  $SE = 0.22$ ,  $z(6917) = -13.1$ ,  $p < .001$ ) and 6 channels ( $\beta = -0.66$ ,  $SE = 0.19$ ,  $z(6917) = -3.42$ ,  $p < .001$ ). There was a significant main effect of target word predictability suggesting that response accuracy was lower at low predictability sentences than both high predictability sentences ( $\beta = 2.18$ ,  $SE = 0.3$ ,  $z(6917) = 7.2$ ,  $p < .001$ ) and medium predictability sentences ( $\beta = -0.52$ ,  $SE = 0.27$ ,  $z(6917) = -1.97$ ,  $p = .049$ ). We also found a significant interaction between channel condition and target word predictability ( $\beta = -0.71$ ,  $SE = 0.29$ ,  $z(6917) = -2.44$ ,  $p = .015$ ).

We performed a subsequent subgroup analyses on each noise channel condition. They revealed that the interaction was driven by the effect of predictability at 4 channels: The accuracy at high predictability sentences was higher than medium predictability sentences ( $\beta = 1.14$ ,  $SE = 0.37$ ,  $z(1608) = 3.1$ ,  $p < .001$ ), which in turn was also higher than low predictability sentences ( $\beta = 1$ ,  $SE = 0.24$ ,  $z(1608) = 4.2$ ,  $p < .001$ ). There was no significant difference in response accuracy between low predictability and high predictability sentences at both 6 channels ( $\beta = 0.33$ ,  $SE = 0.32$ ,  $z(2590) = 1.04$ ,  $p = .3$ ) and 8 channels ( $\beta = -0.014$ ,  $SE = 0.32$ ,  $z(2719) =$

-0.04,  $p = .97$ ). However, response accuracy was higher in high predictability than in medium predictability sentences at both 6 channels ( $\beta = 1.83$ ,  $SE = 0.65$ ,  $z(2590) = 2.83$ ,  $p < .005$ ) and 8 channels ( $\beta = 1.54$ ,  $SE = 0.61$ ,  $z(2719) = 2.54$ ,  $p = .011$ ).

We also found a significant main effect of global channel context which showed that the response accuracy was higher in predictable channel context than in unpredictable channel context ( $\beta = -0.27$ ,  $SE = 0.14$ ,  $z(6917) = -2.02$ ,  $p = .04$ ).

Further, to test the effect of practice on adaptation to degraded speech, we added trial number as a fixed effect in the maximal model. Note that there were 30 trials in each block in the predictable channel context (i.e., blocked design). For comparability, we divided unpredictable channel context (i.e., randomized design) into four blocks. Then following the same procedure as above, we obtained an optimal model. We did not find a significant main effect of trial number indicating that the response accuracy did not change throughout the experiment ( $\beta = -0.0004$ ,  $SE = 0.01$ ,  $z(6917) = -0.05$ ,  $p = 0.97$ ). It remained constant within each block in the predictable channel context ( $\beta = -0.02$ ,  $SE = 0.01$ ,  $z(3291) = -1.43$ ,  $p = 0.15$ ) as well as in the unpredictable channel context ( $\beta = 0.01$ ,  $SE = 0.01$ ,  $z(3291) = 1.05$ ,  $p = 0.29$ ).

## 1.5 Discussion

The present study had three goals: i) to examine if previously reported facilitatory effect of semantic predictability is restricted to only highly predictable sentence endings; ii) to assess the role of perceptual adaptation on the facilitation of language comprehension by sentence predictability; and iii) to use and establish a sensitive metric to measure language comprehension that takes into account whether listeners benefited from the semantic context of the sentence they have listened to.

Results of our study showed the expected interaction between predictability and degraded speech, that is, language comprehension was better for high-cloze than for low-cloze target words when the speech signal was moderately degraded by noise-vocoding through 4 channels, while the effect of predictability was absent when

speech was not intelligible (noise-vocoding through 1 channel). These results are fully in line with Obleser and Kotz [8]; we partly included identical sentences from their study in the present study (see Appendix A). Importantly, in contrast to their study, we had also created sentences with medium-cloze target words (which were intermediate between high-cloze and low-cloze target words) and found that the effect of predictability was also significant when comparing sentences with medium-cloze target words against sentences with low-cloze target words at 4 channels noise-vocoding condition. Recognition of a target word was dependent on its level of predictability (measured by cloze probability), and correct recognition was not just limited to high-cloze target words. These significant differences in response accuracy between medium-cloze and low-cloze target words, and between medium-cloze and high-cloze target words at noise-vocoding through 4 channels show that the sentence-final word recognition is facilitated by semantic predictability in a graded manner. This is in line with the findings from the ERP literature where it has been observed that semantic predictability, in terms of cloze probability of target word of a sentence, modulates semantic processing, indexed by N400, in a graded manner [1, 2, 33].

The interpretation of the observed graded effect of semantic predictability at the moderate level of spectral degradation (i.e., at noise-vocoding through 4 channels) provides a novel insight into how listeners form prediction when the bottom-up input is compromised. That is, in an adverse listening condition, listeners rely more on top-down semantic prediction than on bottom-up acoustic-phonetic cues. However, such a reliance on top-down prediction is not an all-or-none phenomenon; instead, listeners form a probabilistic prediction of the target word. The effect of target word predictability on comprehension is not sharply focused solely on high-cloze target words like a ‘searchlight’. But rather it is spread across a wide range including low-cloze and medium-cloze target words. As the cloze probability of the target words decreases from high to low, the focus of the searchlight becomes less precise.

## 1.6 Conclusion

In conclusion, this study provides novel insights into predictive language processing when bottom-up signal quality is compromised and uncertain: We show that while processing moderately degraded speech, listeners form top-down predictions across a wide range of semantic space that is not restricted within highly predictable sentence endings. In contrast to the narrowed expectation view, comprehension of words ranging from low- to high-cloze probability, including medium-cloze probability, is facilitated in a graded manner while listening to a moderately degraded speech. We also found better speech comprehension when individuals were likely to have adapted to the noise condition in the blocked design compared to the randomized design. We did not find learning effects at the trial-to-trial level of perceptual adaption – it may be that the adaptation was hampered by variation in higher-level semantic features (i.e., target word predictability). We also argue that for the examination of semantic predictability effects during language comprehension, the analyses of response accuracy should be based on the trials in which context evoking words are correctly identified in the first place to make sure that listeners make use of the contextual cues instead of analyzing general word recognition scores.

# References

- [1] Mante S Nieuwland et al. “Large-scale replication study reveals a limit on probabilistic prediction in language comprehension”. In: *eLife* 7 (2018), pp. 1–24.
- [2] Katherine A DeLong, Thomas P Urbach, and Marta Kutas. “Probabilistic word pre-activation during language comprehension inferred from electrical brain activity”. In: *Nature Neuroscience* 8.8 (July 10, 2005), pp. 1117–1121. URL: <http://dx.doi.org/10.1038/nn1504>.
- [3] Micha Heilbron et al. “A hierarchy of linguistic predictions during natural language comprehension”. In: *bioRxiv* (2020).
- [4] Antje Strauß, Sonja A. Kotz, and Jonas Obleser. “Narrowed Expectancies under Degraded Speech: Revisiting the N400”. In: *Journal of Cognitive Neuroscience* 25.8 (Aug. 2013), pp. 1383–1395. URL: [http://dx.doi.org/10.1162/jocn\\_a\\_00389](http://dx.doi.org/10.1162/jocn_a_00389).
- [5] Signy Sheldon, M. Kathleen Pichora-Fuller, and Bruce A. Schneider. “Priming and sentence context support listening to noise-vocoded speech by younger and older adults”. In: *The Journal of the Acoustical Society of America* 123.1 (Feb. 2008), pp. 489–499. URL: <http://dx.doi.org/10.1121/1.2783762>.
- [6] Ruth E. Corps and Hugh Rabagliati. “How top-down processing enhances comprehension of noise-vocoded speech: Predictions about meaning are more important than predictions about form”. In: *Journal of Memory and Language* 113 (Aug. 2020), p. 104114. URL: <http://dx.doi.org/10.1016/j.jml.2020.104114>.
- [7] Jonas Obleser et al. “Functional Integration across Brain Regions Improves Speech Perception under Adverse Listening Conditions”. In: *Journal of Neuroscience* 27.9 (Feb. 28, 2007), pp. 2283–2289. URL: <http://dx.doi.org/10.1523/JNEUROSCI.4663-06.2007>.
- [8] Jonas Obleser and S. A. Kotz. “Expectancy Constraints in Degraded Speech Modulate the Language Comprehension Network”. In: *Cerebral Cortex* 20.3 (June 26, 2009), pp. 633–640. URL: <http://dx.doi.org/10.1093/cercor/bhp128>.
- [9] Jonas Obleser and Sonja A. Kotz. “Multiple brain signatures of integration in the comprehension of degraded speech”. In: *NeuroImage* 55.2 (2011), pp. 713–723. URL: <https://www.sciencedirect.com/science/article/pii/S1053811910016034>.
- [10] Jonas Obleser. “Putting the Listening Brain in Context”. In: *Language and Linguistics Compass* 8.12 (Dec. 2014), pp. 646–658. URL: <http://dx.doi.org/10.1111/lnc3.12098>.
- [11] Emmanuel Dupoux and Kerry Green. “Perceptual adjustment to highly compressed speech: Effects of talker and rate changes.” In: *Journal of Experimental Psychology: Human Perception and Performance* 23.3 (1997), pp. 914–927. URL: <http://dx.doi.org/10.1037/0096-1523.23.3.914>.

- [12] Arthur G. Samuel and Tanya Kraljic. “Perceptual learning for speech”. In: *Attention, Perception, & Psychophysics* 71.6 (Aug. 2009), pp. 1207–1218. URL: <http://dx.doi.org/10.3758/APP.71.6.1207>.
- [13] Sara Guediche et al. “Speech perception under adverse conditions: Insights from behavioral, computational, and neuroscience research”. In: *Frontiers in Systems Neuroscience* 7.Jan (2014), pp. 1–16.
- [14] Stuart Rosen, Andrew Faulkner, and Lucy Wilkinson. “Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants”. In: *The Journal of the Acoustical Society of America* 106.6 (Dec. 1999), pp. 3629–3636. URL: <http://dx.doi.org/10.1121/1.428215>.
- [15] Matthew H. Davis et al. “Lexical Information Drives Perceptual Learning of Distorted Speech: Evidence From the Comprehension of Noise-Vocoded Sentences.” In: *Journal of Experimental Psychology: General* 134.2 (2005), pp. 222–241. URL: <http://dx.doi.org/10.1037/0096-3445.134.2.222>.
- [16] J. Erb et al. “The Brain Dynamics of Rapid Perceptual Adaptation to Adverse Listening Conditions”. In: *Journal of Neuroscience* 33.26 (June 26, 2013), pp. 10688–10697. URL: <http://dx.doi.org/10.1523/JNEUROSCI.4596-12.2013>.
- [17] Sven L. Mattys et al. “Speech recognition in adverse conditions: A review”. In: *Language and Cognitive Processes* 27.7-8 (2012), pp. 953–978.
- [18] Mitchell S. Sommers, Lynne C. Nygaard, and David B. Pisoni. “Stimulus variability and spoken word recognition. I. Effects of variability in speaking rate and overall amplitude”. In: *The Journal of the Acoustical Society of America* 96.3 (Sept. 1994), pp. 1314–1324. URL: <http://dx.doi.org/10.1121/1.411453>.
- [19] Delphine Dahan and James S. Magnuson. “Spoken Word Recognition”. In: Elsevier, 2006, pp. 249–283. URL: <http://dx.doi.org/10.1016/b978-012369374-7/50009-2>.
- [20] Mercedes Atienza, Jose L. Cantero, and Elena Dominguez-Marin. “The time course of neural changes underlying auditory perceptual learning”. In: *Learning and Memory* 9.3 (2002), pp. 138–150.
- [21] Clarissa J. Whitmire and Garrett B. Stanley. “Rapid Sensory Adaptation Redux: A Circuit Perspective”. In: *Neuron* 92.2 (2016), pp. 298–315. URL: <http://dx.doi.org/10.1016/j.neuron.2016.09.046>.
- [22] Kenneth I. Vaden et al. “The cingulo-opercular network provides word-recognition benefit”. In: *Journal of Neuroscience* 33.48 (2013), pp. 18979–18986.
- [23] Kenneth I. Vaden et al. “Cortical activity predicts which older adults recognize speech in noise and when”. In: *Journal of Neuroscience* 35.9 (2015), pp. 3929–3937.
- [24] Kenneth I. Vaden et al. “Cingulo-Opercular Function During Word Recognition in Noise for Older Adults with Hearing Loss”. In: *Experimental Aging Research* 42.1 (Dec. 18, 2015), pp. 67–82. URL: <http://dx.doi.org/10.1080/0361073X.2016.1108784>.
- [25] Mark A. Eckert, Susan Teubner-Rhodes, and Kenneth I. Vaden. “Is Listening in Noise Worth It? The Neurobiology of Speech Recognition in Challenging Listening Conditions”. In: *Ear & Hearing* 37.1 (July 2016), 101S–110S. URL: <http://dx.doi.org/10.1097/aud.0000000000000300>.

- [26] Gesa Hartwigsen, Thomas Golombek, and Jonas Obleser. “Repetitive transcranial magnetic stimulation over left angular gyrus modulates the predictability gain in degraded speech comprehension”. In: *Cortex* 68 (July 2015), pp. 100–110. URL: <http://dx.doi.org/10.1016/j.cortex.2014.08.027>.
- [27] Douglas Bates et al. “Fitting Linear Mixed-Effects Models Using lme4”. In: *Journal of Statistical Software* 67.1 (2015). URL: <http://dx.doi.org/10.18637/jss.v067.i01>.
- [28] Alexandra Kuznetsova, Per B. Brockhoff, and Rune H. B. Christensen. “lmerTest Package: Tests in Linear Mixed Effects Models”. In: *Journal of Statistical Software* 82.13 (2017). URL: <http://dx.doi.org/10.18637/jss.v082.i13>.
- [29] Dale J. Barr et al. “Random effects structure for confirmatory hypothesis testing: Keep it maximal”. In: *Journal of Memory and Language* 68.3 (Apr. 2013), pp. 255–278. URL: <http://dx.doi.org/10.1016/j.jml.2012.11.001>.
- [30] C. E. Grueber et al. “Multimodel inference in ecology and evolution: challenges and solutions”. In: *Journal of Evolutionary Biology* 24.4 (Jan. 27, 2011), pp. 699–711. URL: <http://dx.doi.org/10.1111/j.1420-9101.2010.02210.x>.
- [31] Shane A. Richards, Mark J. Whittingham, and Philip A. Stephens. “Model selection and model averaging in behavioural ecology: the utility of the IT-AIC framework”. In: *Behavioral Ecology and Sociobiology* 65.1 (Aug. 2011), pp. 77–89. URL: <http://dx.doi.org/10.1007/s00265-010-1035-8>.
- [32] Douglas Bates et al. “Parsimonious Mixed Models”. In: *arXiv* (2015). arXiv: 1506.04967. URL: <http://arxiv.org/abs/1506.04967>.
- [33] Edward W. Wlotko and Kara D. Federmeier. “Age-related changes in the impact of contextual strength on multiple aspects of sentence comprehension”. In: *Psychophysiology* 49.6 (Apr. 3, 2012), pp. 770–785. URL: <http://dx.doi.org/10.1111/j.1469-8986.2012.01366.x>.