

Individual differences in selective attention predict speech identification at a cocktail party

Daniel Oberfeld*, Felicitas Klöckner-Nowotny

Department of Psychology, Section Experimental Psychology, Johannes Gutenberg-Universität, Mainz, Germany

Abstract Listeners with normal hearing show considerable individual differences in speech understanding when competing speakers are present, as in a crowded restaurant. Here, we show that one source of this variance are individual differences in the ability to focus selective attention on a target stimulus in the presence of distractors. In 50 young normal-hearing listeners, the performance in tasks measuring auditory and visual selective attention was associated with sentence identification in the presence of spatially separated competing speakers. Together, the measures of selective attention explained a similar proportion of variance as the binaural sensitivity for the acoustic temporal fine structure. Working memory span, age, and audiometric thresholds showed no significant association with speech understanding. These results suggest that a reduced ability to focus attention on a target is one reason why some listeners with normal hearing sensitivity have difficulty communicating in situations with background noise.

DOI: [10.7554/eLife.16747.001](https://doi.org/10.7554/eLife.16747.001)

Introduction

Imagine yourself sitting at a table in a crowded restaurant, chatting with a friend of yours. Can you follow the conversation with your friend easily, or does the high noise level in general or more specifically the conversations heard from other tables interfere with speech intelligibility? While the human auditory system has impressive abilities in structuring the mixture of sound waves arriving at the ears into different auditory objects or streams (auditory scene analysis; e.g., *Bregman, 1990; Carlyon, 2004*), listeners show considerable variation when it comes to speech understanding in adverse acoustic conditions, such as the almost proverbial 'cocktail-party' situation described above (*Bronkhorst, 2000; Cherry, 1953*). Surprisingly, pronounced individual differences in speech identification in background noise are observed even in listeners with *normal hearing*, that is, with audiometric thresholds better than 20 dB HL within the frequency range most important for speech (100 Hz–4 kHz; *Byrne et al., 1994*), usually taken as an indication of approximately intact processing in the inner ear. For example, *Ruggles and Shinn-Cunningham (2011)* tested normal-hearing subjects in a simulated cocktail-party listening task with two competing speakers presented 15° to the left and right of the target speaker. Across listeners, the percentage of correct responses in reporting a sequence of digits produced by the target speaker varied between 40% and 85% in an anechoic condition. Compatible with these experimental results, in clinical settings a relevant number of patients with normal audiometric findings complain about hearing difficulties in daily life (*Zhao and Stephens, 2007*).

Understanding speech in a cocktail-party situation with interfering speakers and other background noise requires *selective attention* (*Bronkhorst, 2015; Cherry, 1953; Shinn-Cunningham, 2008; Xiang et al., 2010*), and was even proposed to be the 'best-known real life example of *selective attention*' (*Pashler, 1998*). The information from the target speaker needs to be processed,

*For correspondence: oberfeld@uni-mainz.de

Competing interests: The authors declare that no competing interests exist.

Funding: See page 19

Received: 07 April 2016

Accepted: 08 August 2016

Published: 31 August 2016

Reviewing editor: Barbara G Shinn-Cunningham, Boston University, United States

© Copyright Oberfeld and Klöckner-Nowotny. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

while information from other sound sources should be ignored. The main hypothesis tested in the present study was that individual differences in the *capability to direct auditory selective attention to the relevant stimulus in the presence of distractors* explain a significant proportion of the inter-individual variance in cocktail-party listening performance. Although several aspects of speech appear to be processed outside the focus of attention (*Pulvermüller and Shtyrov, 2006*), attention enhances the representation of speech (and other sounds) at relatively early stages (e.g., *Choi et al., 2014; Srinivasan et al., 2012; Xiang et al., 2010; Zion Golombic et al., 2013*). On the behavioral level, the importance of attention is illustrated by studies that manipulated the a-priori information concerning the target speaker. For instance, in an experiment by *Kidd et al. (2005)*; the speech identification performance was better when the listener knew in advance which of three talkers (presented at different spatial positions) would be the target speaker, compared to conditions where the target location was uncertain. Thus, for exactly identical acoustic signals, being able to direct attention to the correct location results in a large improvement in speech recognition (e.g., *Best et al., 2007; Kitterick et al., 2010*). The direction of selective attention to the target speaker can be impaired due to limitations imposed by the *acoustic signal* or by *perceptual or cognitive characteristics of the listener* (cf. *Mattys et al., 2012*). If a listener has problems in using acoustic cues for the formation of auditory objects or streams, or on a more cognitive level is incapable of ignoring irrelevant information, then speech identification performance will be low even when fundamental frequency, timbre, and spatial location of the speakers differ.

With respect to listeners' attentional capabilities, it is important to take into consideration that attention is a multifaceted phenomenon (cf. *Styles, 2006*). In a cocktail party situation, it is required to attend to a certain speaker (target) and to ignore the other sound sources (distractors). We were interested in whether speech understanding in a cocktail-party situation could be related to a more general ability to focus attention on a target in the presence of distractors. Surprisingly, this particular aspect of attention has not been studied very systematically in previous experiments that investigated the role of attention for speech understanding in noise (*Füllgrabe et al., 2014; Gatehouse and Akeroyd, 2008; Heinrich et al., 2015; Neher et al., 2009, 2011, 2012; Schoof and Rosen, 2014; van Rooij et al., 1989*). Most of these studies used tests developed for neuropsychological settings, like the Test of Everyday Attention (TEA; *Robertson et al., 1996*) and the Trail Making Tests (TMT; *Bowie and Harvey, 2006*), indexing *visual search* (TEA Map Search, Telephone Search), *task/attentional switching* (TEA Visual Elevator and Auditory Elevator with Reversal, TMT-B), *sustained attention* (TMT-A, TEA Lottery Test and Elevator Counting), or *divided attention* (TEA Telephone Search While Counting). Only one TEA subtest ("Elevator counting with distraction") directly addresses the capability to ignore distractors. Here, participants have to count low pitch tones ('targets') while ignoring interspersed high pitch tones ('distractors'). We are aware of only two studies that included this subtest (*Gatehouse and Akeroyd, 2008; Neher et al., 2009*), in hearing-impaired listeners. Concerning the other tasks, while switching attention is relevant for situations where the target speaker changes dynamically, for example in a conversation involving more than two persons, there are many situations where the target speaker does not change. Visual search and sustained attention seem even less relevant for cocktail-party listening. For this reason, our study included tasks in which subjects had to *identify an auditory or visual target element in the presence of distractors*. In our view, this is the most important aspect of attention in cocktail-party listening.

As a measure of *visual selective attention*, we used a *flanker task* as established by *Eriksen and Eriksen (1974)*; where a target stimulus is surrounded by task-irrelevant distractors (flankers). In the critical *incompatible condition*, the flankers and the target call for opposite responses. If the incompatible flankers produce only small response time (RT) costs, then the participant has a high ability to focus visual selective attention on the target stimulus. This flanker interference is defined as the difference between the average response time (RT) in the incompatible condition and in a neutral condition where the flankers are not associated with one of the responses relevant for the target.

To measure the individual ability to *direct auditory selective attention* to a target stimulus while ignoring distractors, we used an intensity discrimination task under backward masking. If – as in the present study – a target sound is followed by a backward masker after a silent inter-stimulus interval (ISI) of 50 ms or more, it is virtually impossible that the masker affects the representation of the target in the auditory nerve (*Kiang et al., 1965; Plack and Viemeister, 1992*). Instead, strong effects of the backward masker on intensity discrimination can be explained by a failure to selectively attend

to the target sounds while ignoring the maskers (*Oberfeld and Stahn, 2012; Oberfeld et al., 2012; Schlauch et al., 1997*). For example, in a study from our lab that quantified the amount of attention directed to the maskers using a behavioral reverse-correlation approach (*Oberfeld et al., 2014*), the effect of non-simultaneous masking on the intensity difference limen (DL) was well accounted for by the attention to the maskers, explaining 72% of the variance.

Listeners with normal audiometric thresholds may differ in their sensitivity to the temporal fine structure (TFS) of sounds (e.g., *Füllgrabe, 2013; Ruggles et al., 2011*), which is necessary for using interaural time difference (ITD) cues to sound localization. These differences were proposed to be due to cochlear neuropathy, which could for instance be caused by moderate noise exposure (*Bharadwaj et al., 2015; Kujawa and Liberman, 2009*) and is sometimes described as 'hidden hearing loss' (*Plack et al., 2014*) because it cannot be detected using standard measures of audiometric threshold. Several studies showed a correlation between TFS sensitivity and the recognition of speech in noise, for normal-hearing as well as for hearing-impaired listeners (*Bharadwaj et al., 2015; Füllgrabe et al., 2014; Neher et al., 2011, 2012; Ruggles et al., 2011; Schoof and Rosen, 2014*). For this reason, our study included binaural sensitivity for the temporal fine structure as a potential predictor of speech identification in a cocktail-party situation, using a task proposed by *Hopkins and Moore (2010)* that measures the smallest detectable interaural phase difference (IPD) of a sinusoidal carrier relative to an IPD of 0°.

As additional cognitive measures, working memory capacity (e.g., *Akeroyd, 2008; Füllgrabe and Rosen, 2016*) measured in a sentence span test (*Daneman and Carpenter, 1980*), and processing speed (e.g., *Salthouse, 1996; Tun and Wingfield, 1999*) measured by the RT in the neutral condition of the visual flanker task, were included as potential predictors of speech-in-noise identification. The latter was measured in a simulated cocktail-party listening situation with two competing speakers that were presented 25° to the left and right of the target speaker, who was positioned in front of the listener (azimuthal angle 0°). In addition, self-reported hearing-related problems in daily life were assessed via the Speech, Spatial and Qualities of Hearing Scale (SSQ) by *Gatehouse and Noble (2004)*, using the German version (*Kießling et al., 2011*).

Results

To which extent did speech understanding in a cocktail-party situation depend on the capability of directing selective attention to a target in the presence of distractors, binaural sensitivity for the temporal fine structure (TFS), and other factors? To answer this question, a multiple linear regression analysis was conducted (the statistical details are described in Materials and methods). The criterion variable was the speech recognition score (SRS) defined as the proportion correct in the simulated cocktail-party listening task with two interfering speakers. We used a sentence identification task based on a German matrix test (*Wagener et al., 1999a*) and presented binaural simulations of an anechoic environment (see Materials and methods). The nine predictors were (1) the elevation of the intensity difference limen caused by the backward masker (DL_{elev}), defined as the difference between the DL under masking and the DL in quiet, which measures the capability of directing auditory selective attention to a target (*Oberfeld et al., 2014*), (2) the amount of flanker interference in the flanker task (Int_{Flanker}), which indexes visual selective attention, (3) the IPD threshold in the TFS-LF (*Hopkins and Moore, 2010*) test (TFS_{th}), which measures binaural sensitivity for the temporal fine structure, (4) the pure-tone average threshold on the better ear (PTA_{BE}) at octave frequencies between 125 Hz and 4 kHz, (5) the average asymmetry in the hearing thresholds between left and right ear in the same frequency range (HL_{diff}), (6) the intensity-DL in quiet (DL_{quiet}), which represents a suprathreshold measure of hearing ability that is not related to selective attention, (7) the response time in the neutral condition of the flanker task (RT_{neutral}), which was included as a measure of processing speed (*Salthouse, 2000*), and (8) the proportion of correctly recalled consonants in the sentence span task (SS_{Pcorr}) that indexes working memory capacity. Finally, (9) the age of the participant was added as a predictor, as in previous studies (e.g., *Neher et al., 2012*), to investigate whether the observed inter-individual differences in cocktail-party listening are determined by other factors related to age. Note that due to the relatively large sample size it was not necessary to summarize the different predictors into a small number of factors as in some previous studies (*Füllgrabe et al., 2014; Heinrich et al., 2015; Schoof and Rosen, 2014; van Rooij et al., 1989*).

Table 1. Results of the multiple regression analysis. Criterion variable: speech recognition score (SRS; proportion correct) in the simulated cocktail-party listening task. Predictors: age, masker-induced elevation of the intensity difference limen (DL_{elev}), the amount of flanker interference in the flanker task ($Int_{Flanker}$), IPD threshold in the TFS-LF task (TFS_{th}), pure-tone average thresholds on the better ear (PTA_{BE}), average asymmetry in the hearing thresholds between left and right ear (HL_{diff}), intensity-DL in quiet (DL_{quiet}), response time in the neutral condition of the flanker task ($RT_{neutral}$), and proportion of correctly recalled consonants in the working memory task (SS_{Pcorr}). All variables were z-standardized.

| Predictor | β | SE | t | p | GDW | β_{Lasso} |
|-----------------|---------------|--------------|--------------|--------------|-----------------------------|-----------------|
| Intercept | 0.068 | 0.096 | 0.710 | 0.480 | | 0.081 |
| Age | 0.194 | 0.110 | 1.760 | 0.086 | 0.020 | – |
| DL_{elev} | –0.347 | 0.107 | 3.240 | 0.003 | 0.152 | –0.220 |
| $Int_{Flanker}$ | –0.233 | 0.103 | 2.270 | 0.029 | 0.052 | –0.081 |
| TFS_{th} | –0.383 | 0.103 | 3.730 | 0.001 | 0.204 | –0.286 |
| PTA_{BE} | 0.137 | 0.102 | 1.350 | 0.186 | 0.016 | – |
| HL_{diff} | –0.088 | 0.106 | 0.830 | 0.413 | 0.007 | – |
| DL_{quiet} | –0.007 | 0.114 | 0.070 | 0.948 | 0.021 | |
| $RT_{neutral}$ | –0.037 | 0.129 | 0.280 | 0.778 | 0.015 | – |
| SS_{Pcorr} | 0.193 | 0.111 | 1.740 | 0.091 | 0.085 | 0.089 |
| | | | | | $R^2 = 0.57$ $p < 0.001$ | $R^2 = 0.44$ |

Note: $N = 45$. β : estimated ordinary least-squares (OLS) regression coefficient. SE: standard error of the estimate. t: t-statistic. Bold font indicates a β significantly different from 0 ($p < 0.05$). GDW: general dominance weight. β_{Lasso} : regression coefficients for predictors selected by the Lasso procedure (model selection via four-fold cross-validation).

DOI: [10.7554/eLife.16747.002](https://doi.org/10.7554/eLife.16747.002)

The regression model showed a good fit, $R^2 = 0.57$, $p < 0.001$, $N = 45$. As can be seen in **Table 1**, the performance in the cocktail-party listening task was significantly negatively related to the intensity-DL elevation under backward masking. Thus, compatible with our hypotheses, participants who showed a better capability of focusing attention on the target sounds in the intensity discrimination task were less affected by the interfering speakers in the cocktail-party listening task. In the same line of reasoning, the significant negative regression coefficient for flanker interference shows that a high capability of directing visual selective attention corresponded to good performance on the cocktail-party listening task. The IPD threshold measured in the TFS-LF task was also significantly negatively related to the SRS. Thus, compatible with previous studies (e.g., [Füllgrabe et al., 2014](#); [Neher et al., 2011, 2012](#); [Ruggles et al., 2012](#)), listeners who showed high sensitivity for the TFS performed better in the spatial listening task. None of the remaining predictors showed a significant association with the performance in the spatial listening task. Notably, neither for age nor for working memory capacity did the regression coefficient differ significantly from 0.

What can be concluded about the relative importance of the different psychoacoustic and cognitive predictors for explaining individual differences in cocktail-party listening? In our data, the nine predictors were partly correlated (see **Table 2**). In such a case, it can be misleading to gauge the relative importance of the predictors by considering the squared standardized regression coefficients (cf. [Tonidandel and LeBreton, 2011](#)). For this reason, we used the ‘dominance analysis’ approach proposed by [Budescu \(1993\)](#); which was shown to be a useful measure of the relative importance of predictors in a regression model, both on theoretical grounds and in simulation studies ([LeBreton et al., 2004](#); [Thomas et al., 2014](#); [Tonidandel and LeBreton, 2011](#)). Dominance analysis provides a quantitative measure of relative importance by examining the change in the variance-accounted-for (ΔR^2) resulting from adding a predictor to all possible regression models containing subsets of the predictors. For example, if there are three predictors (A, B, and C), then there are four possible subset models to which predictor C can be added (that is, models containing only the

Table 2. Pairwise Pearson partial correlation coefficients, controlling for age. $N = 50$. In each row, the upper numbers are the partial correlation coefficients (ρ_{partial}), and the lower numbers are the p -values for the test of $|\rho_{\text{partial}}| > 0$. The rightmost column shows Pearson correlation coefficients with age. Bold font: $p < 0.05$. Italics: $p < 0.10$.

| | DL _{elev} | Int _{Flanker} | TFS _{th} | PTA _{BE} | HL _{diff} | DL _{quiet} | RT _{neutral} | SS _{Pcorr} | SSQ _{speech} | SSQ _{spatial} | SSQ _{qualities} | Age | |
|--------------------------|------------------------|------------------------|------------------------|-------------------|------------------------|------------------------|-----------------------|------------------------|------------------------|------------------------|--------------------------|------------------------|----------------|
| OLSA _{Pcorr} | -0.374 0.008 | -0.149 0.307 | -0.353 0.013 | 0.060 0.683 | -0.163 0.263 | <i>-0.244</i> 0.091 | -0.232 0.109 | 0.338 0.018 | 0.121 0.407 | 0.083 0.570 | 0.230 0.112 | 0.033 0.819 | |
| DL _{elev} | | -0.047 0.748 | 0.038 0.793 | 0.030 0.838 | -0.083 0.570 | -0.045 0.759 | 0.092 0.529 | <i>-0.255</i> 0.077 | -0.302 0.035 | -0.049 0.740 | -0.178 0.220 | -0.045 0.754 | |
| Int _{Flanker} | | | -0.081 0.578 | -0.018 0.903 | -0.234 0.105 | -0.032 0.826 | -0.094 0.522 | -0.112 0.444 | -0.141 0.335 | -0.098 0.501 | -0.112 0.444 | 0.045 0.758 | |
| TFS _{th} | | | | 0.034 0.818 | -0.023 0.873 | 0.399 0.005 | 0.312 0.029 | -0.177 0.224 | -0.149 0.306 | -0.314 0.028 | -0.352 0.013 | 0.027 0.852 | |
| PTA _{BE} | | | | | -0.292 0.042 | -0.083 0.572 | -0.010 0.944 | -0.092 0.531 | 0.136 0.353 | 0.082 0.577 | 0.097 0.508 | <i>-0.248</i> 0.082 | |
| HL _{diff} | | | | | | 0.195 0.180 | 0.038 0.795 | 0.087 0.551 | 0.274 0.057 | 0.198 0.174 | 0.119 0.416 | 0.227 0.113 | |
| DL _{quiet} | | | | | | | 0.383 0.007 | -0.115 0.431 | -0.008 0.959 | -0.118 0.420 | -0.130 0.375 | 0.082 0.573 | |
| RT _{neutral} | | | | | | | | -0.198 0.172 | -0.087 0.552 | -0.135 0.355 | -0.011 0.940 | 0.217 0.129 | |
| SS _{Pcorr} | | | | | | | | | | -0.011 0.940 | 0.081 0.580 | -0.380 0.006 | |
| SSQ _{speech} | | | | | | | | | | | 0.707 <0.0001 | 0.728 <0.0001 | 0.040 0.784 |
| SSQ _{spatial} | | | | | | | | | | | | 0.701 <0.0001 | 0.123 0.393 |
| SSQ _{qualities} | | | | | | | | | | | | | 0.058 0.692 |

DOI: 10.7554/eLife.16747.003

intercept term, intercept and predictor A, intercept and predictor B, and intercept and predictors A and B, respectively). A predictor's *general dominance weight* (GDW; **Azen and Budescu, 2003**) is found by averaging the squared semipartial correlations across all of the possible subset models. This measure indexes a variable's contribution to the prediction of the dependent variable, by itself and in combination with the other predictors. The sum of the GDWs is the total proportion of variance explained by the regression model, R^2 .

As shown in **Table 1**, the general dominance weight was highest for the IPD threshold in the TFS-LF task, followed by the DL-elevation in the intensity discrimination task. According to these results, sensitivity for the binaural TFS and auditory selective attention are the most important predictors of cocktail-party listening. The contribution of flanker interference was much lower, and the GDW for the non-significant predictor sentence span was even slightly higher than the GDW for flanker interference.

To validate the conclusions based on dominance analysis, we used a second approach to variable selection. In the *Lasso* method proposed by **Tibshirani (1996)**; regression coefficients for predictors with only a small explanatory value are set to 0 (*shrinkage*), so that the *Lasso* method effectively performs *subset selection*, that is, selects the most important predictors (**James et al., 2013**). The *Lasso* approach is widely used in the field of statistical learning (**Hastie et al., 2009**). It involves a tuning parameter λ to impose a so-called l_1 -penalty (**Tibshirani, 2011**) on the regression model. We used 4-fold cross-validation for selecting the best model, that is, the optimal value of λ (**James et al.,**

2013). In cross-validation, the set of observations is randomly divided into k groups of approximately equal size. The first group is treated as a validation set, and the model is fit on the remaining $k - 1$ groups. The mean squared error is computed on the observations in the validation set, and this procedure is repeated k times, each time selecting a different group of observations as the validation set. In this approach, first the λ value corresponding to the smallest cross-validation error is selected, and then the regression model is fit to all of the available observations using the selected value of λ . Compatible with the dominance analysis presented above, the model selected by the Lasso method contained the predictors DL_{elev} , TFS_{th} , and Int_{Flanker} (see **Table 1**), and the highest regression coefficients were estimated for DL_{elev} and TFS_{th} . In addition, the Lasso procedure selected the predictor SS_{Pcorr} (see **Table 1**). Thus, unlike the ordinary least-squares (OLS) multiple regression, the Lasso indicated a contribution of working memory span to speech identification in noise.

Finally, it is interesting to compare these multiple regression results to the pairwise partial correlation coefficients with the SRS, controlling for age (see **Table 2**). The predictors DL_{elev} and TFS_{th} showed a significantly negative partial correlation with the SRS, compatible with the results from ordinary least-squares (OLS) regression and the Lasso. As for the Lasso, the partial correlation coefficient for SS_{Pcorr} was also significant. It appears possible that the (moderate) correlations between SS_{Pcorr} and DL_{elev} and age (see **Table 2**) increased the standard error of the regression coefficient for SS_{Pcorr} in the multiple regression analysis shown in **Table 1**. The opposite pattern occurred for Int_{Flanker} . Here, the partial correlation coefficient for Int_{Flanker} was not significant, unlike in the multiple regressions.

The relation between the scores on the SSQ questionnaire (**Gatehouse and Noble, 2004**), representing self-reported hearing abilities in daily life, and the performance in the cocktail-party listening task was analyzed via linear multiple regression. The SRS was the criterion variable, and age and the three SSQ subscales (speech hearing: SSQ_{speech} ; spatial hearing: SSQ_{spatial} ; other qualities: $SSQ_{\text{qualities}}$) were entered as predictors. Using the same criteria as for the regression analysis presented in **Table 1** (see Materials and methods), three participants were excluded as outliers. The model explained only a small, non-significant portion of the variance, $R^2 = 0.137$, $p=0.17$, $N = 47$. Only the regression coefficient for the 'Other qualities' scale was significant, showing a positive relation between this SSQ subscore and the SRS (see **Table 3**). Thus, persons reporting better hearing abilities on the SSQ 'Other qualities' scale tended to perform better in the spatial listening task. As seen in **Table 2**, the partial correlations controlling for age indicated a significant negative relation between the DL-elevation and SSQ_{speech} , and between the IPD threshold in the TFS-LF test and SSQ_{spatial} and $SSQ_{\text{qualities}}$.

Discussion

In a relatively large sample of young, normal-hearing participants ($N = 50$; age range 18–30 years), we studied the role of perceptual and cognitive factors for speech understanding in a cocktail-party situation with spatially separated interfering speakers. Our main hypothesis was that individual differences in the ability to direct auditory selective attention to the relevant stimulus, while ignoring distractors, explain a significant proportion of the inter-individual variance in cocktail-party listening performance. To test this hypothesis, we included tasks that assessed auditory and visual selective

Table 3. Multiple regression analysis of the relation between the SSQ scores (**Gatehouse and Noble, 2004**) representing self-reported hearing abilities (predictors) and the speech recognition score in the simulated cocktail-party listening task (criterion). $N = 47$.

| Predictor | β | SE | t | p |
|--------------------------|---------|-------|-------|-------|
| Intercept | −1.150 | 0.988 | 1.160 | 0.251 |
| Age | −0.006 | 0.034 | 0.170 | 0.864 |
| SSQ_{speech} | 0.006 | 0.116 | 0.050 | 0.958 |
| SSQ_{spatial} | −0.160 | 0.117 | 1.360 | 0.181 |
| $SSQ_{\text{qualities}}$ | 0.321 | 0.142 | 2.250 | 0.030 |

DOI: [10.7554/eLife.16747.004](https://doi.org/10.7554/eLife.16747.004)

attention in the presence of distractors, using non-speech stimuli. Among the many different aspects of attention (cf. *Styles, 2006*), this ability seems particularly relevant in a cocktail-party situation where it is necessary to selectively attend to the target speaker and to ignore the interfering speakers and other background noise (*Bronkhorst, 2015; Shinn-Cunningham, 2008*).

In our experiment, the individual ability to selectively attend to an auditory target stimulus in the presence of distractors (measured in an intensity discrimination task under backward masking; *Oberfeld et al., 2014*) as well as the ability to attend to a visual target stimulus (measured in a Flanker task; *Kramer and Jacobson, 1991*) explained a significant portion of the variance in sentence identification performance with two interfering talkers. Together, the two measures of selective attention explained approximately the same proportion of variance as the binaural TFS sensitivity, which was the predictor with the highest relative importance (see *Table 1*). These results are compatible with our hypothesis that not only rather basic auditory factors like spatial hearing abilities contribute to individual differences in cocktail-party listening, but that in the cognitive domain a general ability of focusing attention on a relevant target stimulus represents an additional important predictor. One of the few previous studies that measured selective attention in the presence of distractors (*Gatehouse and Akeroyd, 2008*) found that in unaided hearing-impaired listeners (mean age about 66 years) the word recognition performance in static background noise was positively related to performance in the 'Elevator counting with distraction' subtest of the Tests of Everyday Attention. However, the analysis did not control for effects of age. In a recent study (*Cahana-Amitay et al., 2016*), a composite score for 'inhibition' that included the error rate in an incongruent condition of the Stroop color-word test (*Stroop, 1935*) was negatively correlated with sentence-final word recognition (controlling for age) in a group of subjects ranging in age between 55 and 84 years.

Our data also illustrate that it is only of limited value to use a broad, unspecific concept of 'attention' and to study associations between speech-in-noise understanding and aspects of attention that are not necessarily related to the requirements of cocktail-party listening. Thus, it seems more appropriate to recognize that attention has many facets (*Styles, 2006*), to not aggregate across the performance on tasks measuring very different aspects of attention in order to define a general 'attention' factor, and to formulate hypotheses concerning the potential importance of a particular aspect of attention for speech identification in noise. In this line of thinking, we note that in the simulated cocktail-party listening task as well as in the two tasks measuring auditory and visual selective attention, the participants had a-priori knowledge of the spatial or temporal position of the target. For this reason, the direction of attention to the target could be viewed as being endogenous (top-down) rather than exogenous (bottom up) (*Posner, 1980*). Limitations in the ability to focus attention while performing these tasks should thus be related to the dorsal fronto-parietal attentional system in the cortex rather than to the ventral network (cf. *Buschman and Miller, 2007; Corbetta and Shulman, 2002*). To test this hypothesis, future experiments on attentional factors influencing cocktail-party listening could include tasks measuring both the endogenous and exogenous orienting of attention, using for instance temporal or spatial cueing (e.g., *Coull and Nobre, 1998; Posner, 1980*). Alternatively, one could argue that in the intensity discrimination task, the onset of the backward masker elicits a capture of attention away from the target tones (e.g., *Desimone and Duncan, 1995; Jonides and Yantis, 1988; Yantis and Jonides, 1990*). Thus, it might even be necessary to further qualify the description of the particular aspect of attention that is indexed by the intensity discrimination task under backward masking and say that it measures the ability to suppress salient, but task-irrelevant auditory events.

The working memory span did not show a strong relation to cocktail-party listening. Thus, although the sentence identification task we presented requires working memory for storing the sequence of five words, the performance on this task appeared to be more strongly limited by failures of selective attention than by memory aspects. This result is compatible with previous data indicating that working memory capacity and speech identification in noise are associated in older, hearing-impaired participants (see *Akeroyd, 2008*), while in normal hearing subjects this correlation is weaker or even absent (*Füllgrabe and Rosen, 2016*). Thus, measures of selective attention should be included in future studies instead of focusing only on working memory capacity. In general, working memory and attention are not independent (*Cowan et al., 2005*). For instance, *Conway et al. (2001)* studied the probability that in a dichotic listening task participants recognize their own name on the ignored channel (*Moray, 1959*), and found that this probability was higher in participants

with a low WM span. Also, the working memory load affects speech understanding in a cocktail-party setting (*Francis, 2010*). In line with these results, in our data the working memory span showed a marginally significant negative correlation (controlling for age) with the DL-elevation (auditory selective attention), see *Table 2*.

Our results confirm the association between binaural TFS sensitivity for speech identification in a spatial listening task, compatible with earlier studies (*Füllgrabe et al., 2014; Neher et al., 2011; Neher et al., 2012; Ruggles et al., 2011; Schoof and Rosen, 2014*). The significant negative association between the IPD threshold and the SRS could be attributed to a reduced benefit from ITD cues in listeners with impaired binaural TFS sensitivity, although a reduction in TFS sensitivity could be also associated with other perceptual impairments beyond sound source localization (*Moore, 2008*). Spatial cues facilitate auditory streaming/grouping (e.g., *Darwin and Hukin, 1999; David et al., 2015*), selective attention (e.g., *Ihfeldt and Shinn-Cunningham, 2008*), and speech recognition (e.g., *Culling et al., 2004*). The importance of spatial cues was reported to increase with the number of interfering sound sources (*Yost et al., 1996*). Thus, one should expect a smaller influence of TFS sensitivity in situations with only one interfering speaker (compared to two as in the present study). Also, the simulated anechoic environment might have caused an overestimation of the importance of ITD cues. In a typical reverberant environment, spatial cues to sound source segregation are reduced (e.g., *Lavandier and Culling, 2010*). It remains to be shown whether the ability to attend to a target in the presence of distractors plays a stronger role than binaural TFS sensitivity in a reverberant setting. On the other hand, the spatial separation of 25° between target and distractor speakers was larger than in some previous studies that used a separation of only 15° (*Ruggles et al., 2011, 2012, Ruggles and Shinn-Cunningham, 2011*). The corresponding stronger ITD cues in our study might have reduced the importance of the sensory representation of the acoustic stimulus for task performance, which would emphasize the relative importance of central factors. Note that in cocktail-party situations in daily life, the spatial separation between the target speakers and competing speakers will often be even larger than 25°.

The self-report measures of hearing abilities (SSQ scores) showed only weak associations with performance in the cocktail-party listening task (controlling for age). In hearing-impaired listeners, some earlier studies (*Gatehouse and Akeroyd, 2008; Heinrich et al., 2015*) also reported rather small correlations between SSQ scores and speech identification in static or amplitude-modulated background noise, even without controlling for age. *Füllgrabe et al. (2014)* found no correlation between SSQ scores and speech-in-noise perception in audiometrically normal-hearing listeners.

Compared to previous studies on predictors of cocktail-party listening, our experiments introduced several methodological improvements. The sample size in our study was larger than in some previous experiments, so that it was possible to measure the influence of several variables on sentence identification in noise, without first combining predictors into a small number of factors. We also used multiple linear regression analyses, rather than pairwise correlations as in some previous experiments. Predictors like hearing thresholds, age, working memory span and TFS sensitivity are partly correlated (see *Table 2*). Multiple linear regression accounts for correlations among the predictors (*Gauss, 1821*), while interpreting pairwise correlations is extremely difficult for a large set of intercorrelated predictors. Unlike most previous studies, we also report the reliability of the measured variables, which in most cases was acceptable to high. Finally, we used two established approaches for assessing the relative importance of predictors, dominance analysis (*Budescu, 1993*) and Lasso (*Tibshirani, 1996*). The Lasso method was proposed to avoid some of the problems of forward or backward selection in stepwise regression (*Harrell, 2015; James et al., 2013*).

Despite the new findings concerning the role of the ability to direct selective attention to a target stimulus for speech identification in noise and the methodological features of our study, there are of course several limitations. First, we included only a single measure each for auditory and selective attention in the presence of distractors. From a psychometric point of view, it would be desirable to include different paradigms, in order to test whether the results generalize to different tasks indexing selective attention, and to increase the reliability of the measures of selective attention. In the auditory modality, we used intensity discrimination under backward masking as an index of selective attention. In this task, the target sounds have to be selected on the basis of their temporal position within a trial. We had decided against a task measuring the *spatial* direction of attention (e.g., *Sach et al., 2000; Spence and Driver, 1994*) because the performance on such a task depends on abilities of spatial hearing, which we assessed separately in terms of the binaural TFS sensitivity.

Also, the precise perception of the temporal structure of speech is important for intelligibility (*Zion Golumbic et al., 2012*), and non-simultaneous masking can negatively affect speech identification (*Dirks and Bower, 1970*). Still, it would be interesting to study whether the spatial direction of attention shows a similar relation to cocktail-party listening as the temporal direction of attention. Spatial and temporal attention were reported to involve different brain areas (e.g., *Michalka et al., 2015*).

Second, we studied a relatively young group of listeners. It remains to be shown whether attentional abilities play a similar role in normal-hearing older subjects, or in hearing-impaired listeners. Age-related changes in selective attention have been reported (*Zanto and Gazzaley, 2014*), just as for other cognitive skills (e.g., *Salthouse, 1996; Sander et al., 2012*). At the same time, the probability of audiometrically relevant hearing losses as well as of 'hidden hearing losses' (*Plack et al., 2014*) increases with age, partly due to noise exposure across the life span. In fact, the TFS sensitivity shows a gradual deterioration with age in normal-hearing listeners (*Füllgrabe, 2013; Grose and Mamo, 2010; King et al., 2014; Ross et al., 2007*). Thus, the relative importance of psychoacoustic and cognitive predictors of speech understanding in a cocktail-party situation might differ between young, middle-aged and older groups, and future research should address this aspect.

Third, our participants were rather homogeneous in terms of education and (being university students) very likely also in terms of socioeconomic status and cognitive aptitude. Thus, it would be desirable to study potential psychoacoustic and cognitive predictors in a less homogeneous and more representative sample.

Fourth, our simulated cocktail-party listening task presented two spatially separated interfering speakers with the same voice as the target speaker, producing sentences with a fixed syntactical structure, and presented with a spatial separation of only 25° from the target speaker. These characteristics likely rendered the task more difficult than a realistic cocktail-party listening situation. On the other hand, communication situations in daily life often include some relatively static background noise in addition to competing speakers, while in our experiment no background noise was presented. Taken together, we assume realistic communication situations to be somewhat, but not dramatically, less difficult than the simulated cocktail-party listening task we presented. For this reason, although the average speech identification performance should be somewhat better in many realistic situations, we do not assume the latter situations to be so easy that individual differences are strongly reduced. Thus, the association between selective attention in the presence of distractors or binaural TFS-sensitivity and speech-in-noise identification should apply to other cocktail-party situations, although ultimately this is an empirical question that should be tested in future experiments.

In conclusion, the individual ability to focus attention on a target stimulus in the presence of distractors explained a significant portion of the inter-individual variance in cocktail-party listening performance in a relatively young sample of normal-hearing listeners. Previous studies had reported that speech identification in multitalker situations is associated not only to auditory abilities such as binaural TFS sensitivity, but also to cognitive factors, predominantly in older and often hearing-impaired listeners (e.g., *Akeroyd, 2008*) but also in young normal-hearing listeners (*Zekveld et al., 2013*). Our results highlight the importance of studying aspects of attention directly relevant for speech identification in noise, rather than measuring associations with less relevant facets of attention such as visual search or to aggregate across very different aspects of attention in order to define a general 'attention' factor.

Materials and methods

Participants

Fifty listeners with normal hearing participated in the experiment voluntarily. All listeners reported normal hearing and no history of hearing disorders, and normal or corrected-to-normal visually acuity. They had hearing thresholds better than 20 dB HL at octave frequencies between 125 Hz and 4 kHz (that is, in the frequency region most important for speech; *Byrne et al., 1994*), calculated on the basis of equivalent threshold sound pressure levels for the Sennheiser HDA 200 earphones (*Han and Poulsen, 1998*). The maximal asymmetry between left and right ear was 15 dB in the frequency range between 125 Hz and 4 kHz. In the high-frequency range, for all but one listener the hearing thresholds were also better than 20 dB HL at 6 and 8 kHz.

The sample size of $N = 50$ was selected so that in the multiple regression analysis containing nine predictors the power to detect a moderate deviation of a single linear regression coefficient from 0 (partial $R^2 = 0.15$) was $1 - \beta = 0.8$ (two-tailed test), with α set to 0.05. According to G*Power (Faul et al., 2009), the required minimum sample size is 47.

All participants were native speakers of German. Most of them were psychology students at the Johannes Gutenberg – Universität Mainz, they received partial course credit or were paid for their participation. The experiment was conducted according to the principles expressed in the Declaration of Helsinki. All listeners participated voluntarily after providing informed written consent, after the topic of the study and potential risks had been explained to them. They were uninformed about the experimental hypotheses. The study was approved by the ethics committee of the Department of Psychology, Johannes Gutenberg-Universität Mainz.

The participants (39 female, 11 male) ranged in age between 18 and 30 years (mean age 21.5 y, SD = 3.1 y). All held the German general qualification for university entrance (Abitur), and 44 of them were psychology students. Since very good grades in secondary school are required for admission in psychology at German universities, the group can be assumed to have relatively high test intelligence, although we did not conduct an intelligence test (Füllgrabe et al., 2014).

Apparatus

The auditory stimuli were generated digitally, played back via an RME (Haimhausen, Germany) ADI/S digital-to-analog converter ($f_s = 44.1$ kHz, 24-bit resolution), attenuated by a TDT (Alachua, FL) PA5 programmable attenuator, buffered by a TDT HB7 headphone buffer, and presented via Sennheiser (Wedemark, Germany) HDA 200 circumaural headphones calibrated according to IEC 318 (1970). The visual stimuli and task instructions were presented on a 17" TFT computer monitor. The experiment was conducted in a double-walled sound-insulated chamber (IAC Acoustics Germany, Niederkrüchten). Responses were collected via a numeric keypad, a computer keyboard, or a mouse, depending on the task.

Tasks

Audiometric thresholds

Detection thresholds were measured bilaterally using Békésy tracking (Békésy, 1947; Hartmann, 2005) with pulsed 270-ms pure tones including 10-ms \cos^2 on- and off-ramps. The starting frequency was 100 Hz. The frequency increased exponentially from tone to tone, at a rate of 1.4 octaves/minute. For each listener and ear, thresholds were computed as the average sound pressure level in a third-octave band around octave frequencies between 125 Hz and 4 kHz. The average hearing levels are shown in Figure 1. The individual better-ear pure tone average threshold (PTA_{BE})

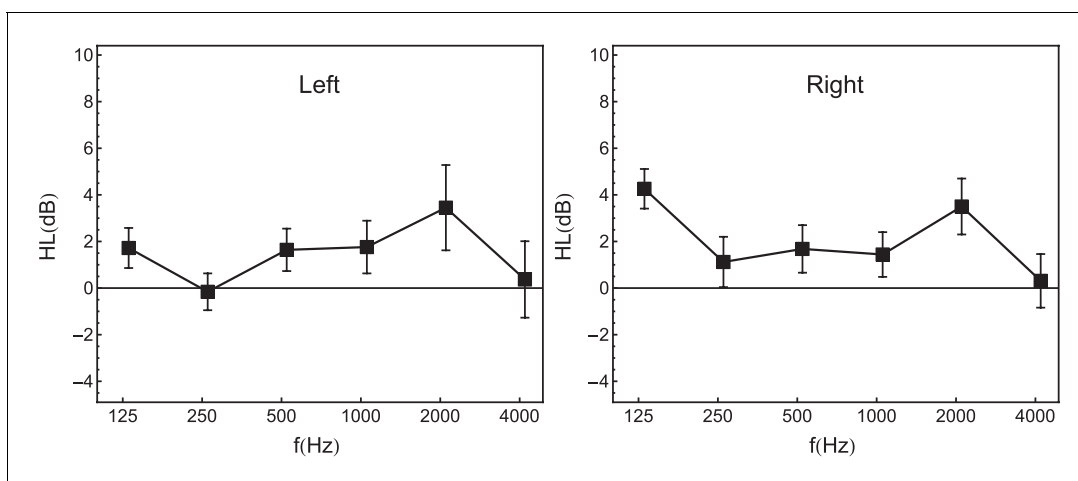


Figure 1. Average audiometric hearing thresholds (in dB HL), at octave frequencies between 125 Hz and 4 kHz ($N = 50$). Left panel: left ear. Right panel: right ear. Error bars represent 95% confidence intervals.

DOI: 10.7554/eLife.16747.005

at octave frequencies between 125 Hz and 4 kHz was entered as a predictor of cocktail-party listening in the regression analyses. In addition, the individual average bilateral asymmetry of the thresholds at the same octave frequencies (HL_{diff}) was included as a predictor, because asymmetric thresholds can affect binaural unmasking on the basis of ITD cues (*Bronkhorst and Plomp, 1989*).

Cocktail-party listening task

As a measure of speech understanding in adverse listening conditions, we used a simulated cocktail-party listening task with two interfering speakers. The listeners performed a sentence identification task with the speech material of the Oldenburger Satztest (OLSA; HörTech gGmbH, Oldenburg), which is a German matrix test (*Wagner et al., 1999a*). The speech material consists of sentences with the syntactic structure *name-verb-numeral-adjective-object* (e.g., 'Peter kauft vier kleine Messer' – 'Peter buys four small knives'). The sentences were constructed by pseudo-randomly selecting one of ten alternatives for each word position. This results in syntactically correct but semantically unpredictable sentences, which makes it possible to use each sentence several times for the same listener. In total, 100 different sentences are available in the OLSA test. The sentences are produced by an adult male speaker and are optimized for similar intelligibility (*Wagner et al., 1999b*). The task was to identify the sentence produced by the target speaker. The matrix of 5 (word position) \times 10 (alternatives) words constituting the sentence test was displayed on a computer monitor. On each trial, subjects were asked to select the five words they had just heard using a computer mouse. The selected words were displayed in a row below the matrix of test words. Initially, the selected words were displayed in black ink. After confirming their selection by clicking on an 'Accept' button, the participants received immediate feedback concerning the correctness of their selection of words. Correctly identified words were colored in green, and incorrect words were colored in red. This visual feedback was presented for 500 ms. The next trial then started automatically after a pause of 500 ms.

The target speaker and the two interfering speakers were presented binaurally via headphones, using head related impulse responses (HRIRs) to simulate the different spatial position of the sound sources. The target speaker was presented from the front (0° azimuthal angle). The interfering speakers were presented 25° to the left and 25° to the right of the target speaker. HRIRs from an anechoic room were used because a previous study showed higher inter-individual differences in speech understanding in an anechoic condition, compared to conditions with reverberation (*Ruggles and Shinn-Cunningham, 2011*). They had been recorded with a head-and-torso simulator Brüel & Kjær Type 4128C at a distance of 80 cm between loudspeaker and microphones and an elevation of 0° (*Kayser et al., 2009*). In the experiment, the target speaker was presented at an average sound pressure level of 58 dB SPL, while each interfering speaker was presented at 60 dB SPL.

Each participant first received five trials without interfering speakers, to become familiar with the task and the response interface. Next, five trials were presented with a single interfering speaker, positioned 25° to the right of the listener. After these brief practice blocks, each listener received three experimental blocks with two interfering speakers (25° left and right), containing 50 trials each. On each trial, the sentences produced by the target speaker and the two interfering speakers were selected at random from the set of 100 test sentences, of course with the restriction that none of the three speakers produced an identical word. Note that the same male voice was used for the target speaker and the two interfering speakers, which made the task relatively difficult (*Cherry, 1953*).

For each listener and each block of 50 trials collected in the sentence identification task with two interfering speakers, the proportion of correctly identified words for the target speaker was computed (speech recognition score; SRS). Because non-normally distributed measures can cause problems in regression/correlation analyses (e.g., *Bishara and Hittner, 2012*) and repeated-measures ANOVAs (e.g., *Oberfeld and Franke, 2013*), the proportions were arcsin-square-root transformed (*Bartlett, 1936*) to obtain a closer approximation to the normal distribution. The data were analyzed with a repeated-measures analysis of variance (rmANOVA), using the multivariate approach. Partial η^2 is reported as a measure of association strength. The same type of rmANOVAs is used in all following analyses. An rmANOVA showed a significant effect of block, $F(2, 48) = 47.34$, $p < 0.001$. The mean proportion of correct responses was considerably lower in the first block than in the two following blocks, compatible with data by *Wagner et al. (1999c)* who reported a sizeable practice effect in steady background noise. For this reason, the data from the first block were excluded from

further analyses. An rmANOVA conducted on the data from blocks 2 and 3 still showed a significant while rather weak effect of block on the SRS, $F(1, 49) = 4.54, p=0.038$, **Cohen (1988)** $d_z = 0.30$. The degree of agreement between the two measurements of the SRS (blocks 2 and 3) represents test-retest reliability and was assessed by an absolute agreement definition of the intraclass correlation in a two-way mixed-model (ICC(A,2) in the nomenclature of **McGraw and Wong, 1996**). The reliability was high, $ICC(A,2) = 0.934$.

Figure 2 shows a histogram of the average individual speech recognition score (proportion correct) in the simulated cocktail-party listening task (blocks 2 and 3). As expected, the listeners showed considerable variation in the SRS. The arcsin-sqrt transformed average individual proportion correct on blocks 2 and 3 (SRS) served as the measure of cocktail-party listening, and was used as the *criterion variable* in the regression analyses.

Auditory intensity discrimination under backward masking (auditory selective attention)

Intensity difference limens (DLs) in quiet and under backward masking were measured using a two-interval, two alternative forced-choice task and an adaptive procedure with a three down, one up rule (**Levitt, 1971**). The targets and the maskers were 1-kHz pure tones with a steady-state duration of 20 ms, presented to the right ear. The tones were gated on and off with 5-ms cosine-squared ramps. The standard level was 60 dB SPL. An intensity increment – that is, a pure tone of the same frequency, duration and temporal envelope – was added in-phase to the standard in one of the observation intervals (selected randomly). The level of the backward masker was 90 dB SPL. The silent interval between standard offset and masker onset was 50 ms (see **Figure 3**). This ISI value is

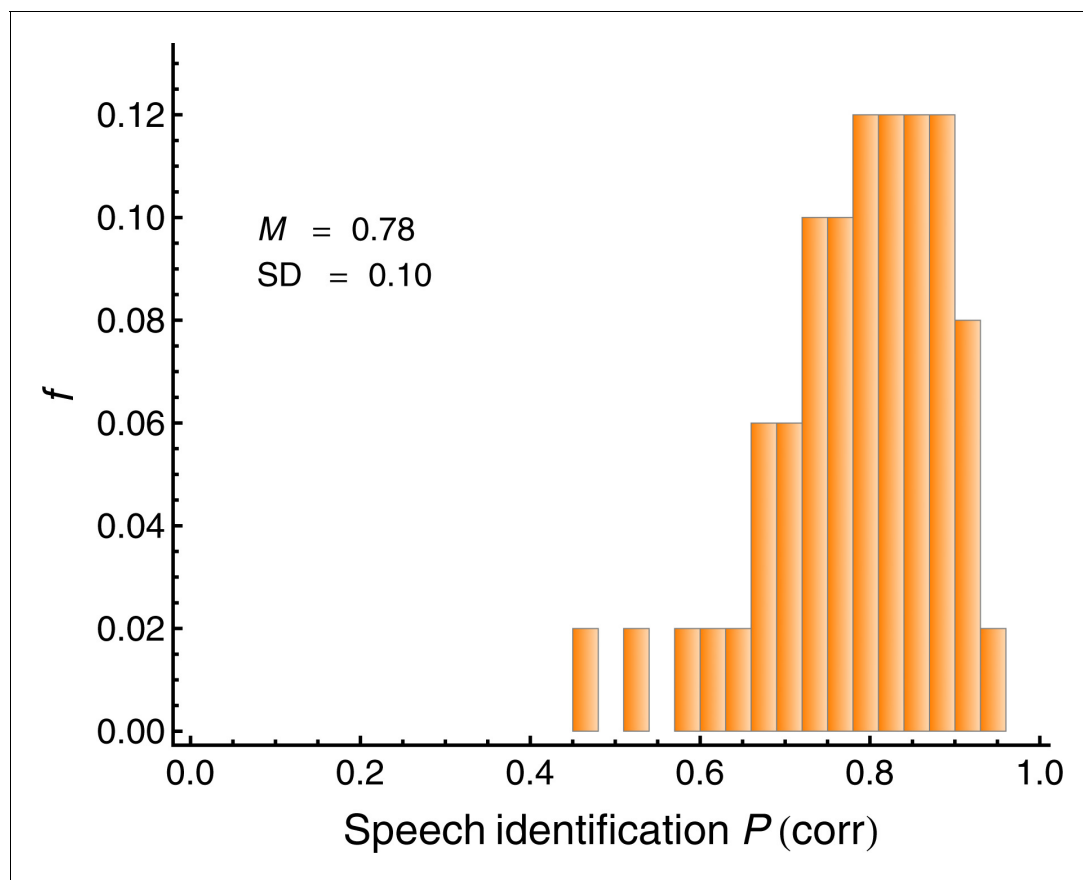


Figure 2. Average individual proportion correct (speech recognition score; SRS) in the simulated cocktail-party listening task with two spatially separated interfering speakers ($N = 50$). This measure served as the criterion variable in the regression analyses. The mean (M) and the standard deviation (SD) are displayed.

DOI: [10.7554/eLife.16747.006](https://doi.org/10.7554/eLife.16747.006)

in the range where the effects of backward and forward masking on speech identification were observed (*Dirks and Bower, 1970*). The temporal interval between the onsets of the two target tones (standard and standard-plus-increment) was 800 ms. The task was to select the interval containing the louder target tone (that is, the standard-plus-increment), and to ignore the maskers. Visual trial-by-trial feedback was provided. In the adaptive procedure, the initial level of the intensity increment, expressed in terms of $10 \log_{10}(\Delta I/I)$, where ΔI is the intensity difference between the standard-plus-increment and the standard and I is the intensity of the standard, was 8 dB. For the in-quiet condition, the step size was 5 dB until the third reversal, and 2 dB for the remaining six reversals. In the backward-masking condition, four reversals were collected with the larger and eight reversals with the smaller step size. The arithmetic mean of $10 \log_{10}(\Delta I/I)$ from the fourth (in quiet) or fifth reversal (backward masking) up to the last even-numbered reversal was taken as the difference limen corresponding to 79.4% correct. Adaptive tracks where the standard deviation of $10 \log_{10}(\Delta I/I)$ at the counting reversals exceeded 7 dB were excluded from the data analysis, which affected 5 tracks (1% of the total of 470 tracks). After a brief practice block, two blocks were obtained in quiet, followed by three blocks under backward masking.

An rmANOVA with the within-subjects factor block (1, 2, 3) showed no significant effect of block on the DL under backward masking, $F(2, 44) = 1.03$, $p=0.90$. Thus, there was no significant practice effect, and therefore the average individual DL under backward masking (DL_{masked}) across the three blocks was computed. The reliability of the masked DL across the three measurements (blocks) was moderate, $ICC(A,3) = 0.871$.

The average DL in quiet (DL_{quiet}) was included as a predictor in the regression analyses, representing a suprathreshold measure of hearing ability that is not related to selective attention. The reliability of DL_{quiet} across the two measurements (blocks) was $ICC(A,2) = 0.660$.

We used the elevation of the intensity-DL caused by the backward masker as a measure of auditory selective attention, as in previous studies (*Oberfeld et al., 2014*). The DL-elevation denotes the difference between the DL under masking and the DL in quiet, $DL_{\text{elev}} = DL_{\text{masked}} - DL_{\text{quiet}}$. As **Figure 4** shows, there was considerable variation in the individual DL-elevations under masking, as expected.

Binaural sensitivity for the temporal fine structure

The binaural sensitivity to temporal fine structure information was measured as the smallest detectable interaural phase difference (IPD) of a sinusoidal carrier relative to an IPD of 0° , using the TFS-LF test proposed by *Hopkins and Moore (2010)*. In a two-interval task, four pure tones (500 Hz) were presented binaurally in each interval. In one of the intervals (selected randomly), the second and fourth tone were presented with an IPD greater than 0° between the right and left ear, while the IPD

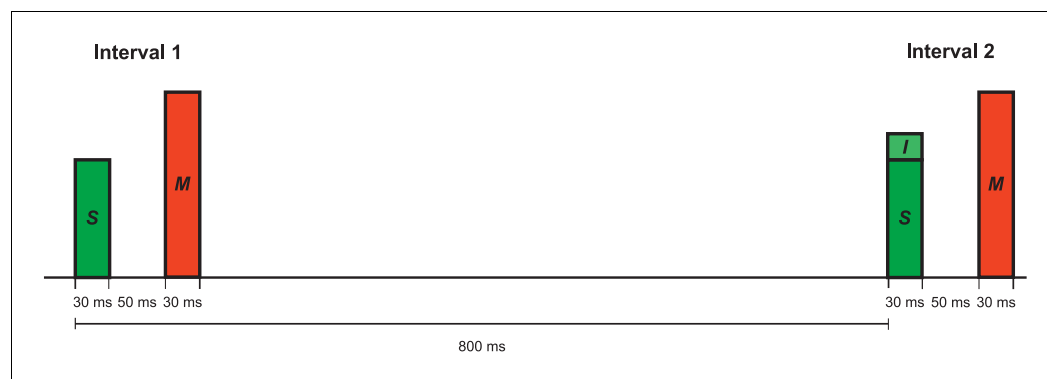


Figure 3. Schematic depiction of the two-interval intensity discrimination task used to measure auditory selective attention. Green: target tones. Red: backward maskers (“distractors”). The standard (S) was a 1 kHz tone presented at 60 dB SPL. An intensity increment (I) was presented in either the first or the second interval, with equal a-priori probability. The task was to select the interval containing the louder target (that is, standard-plus-increment). The maskers were 1 kHz tones presented at 90 dB SPL. The same temporal configuration was used in the in-quiet condition, except that the maskers were not presented.

DOI: [10.7554/eLife.16747.007](https://doi.org/10.7554/eLife.16747.007)

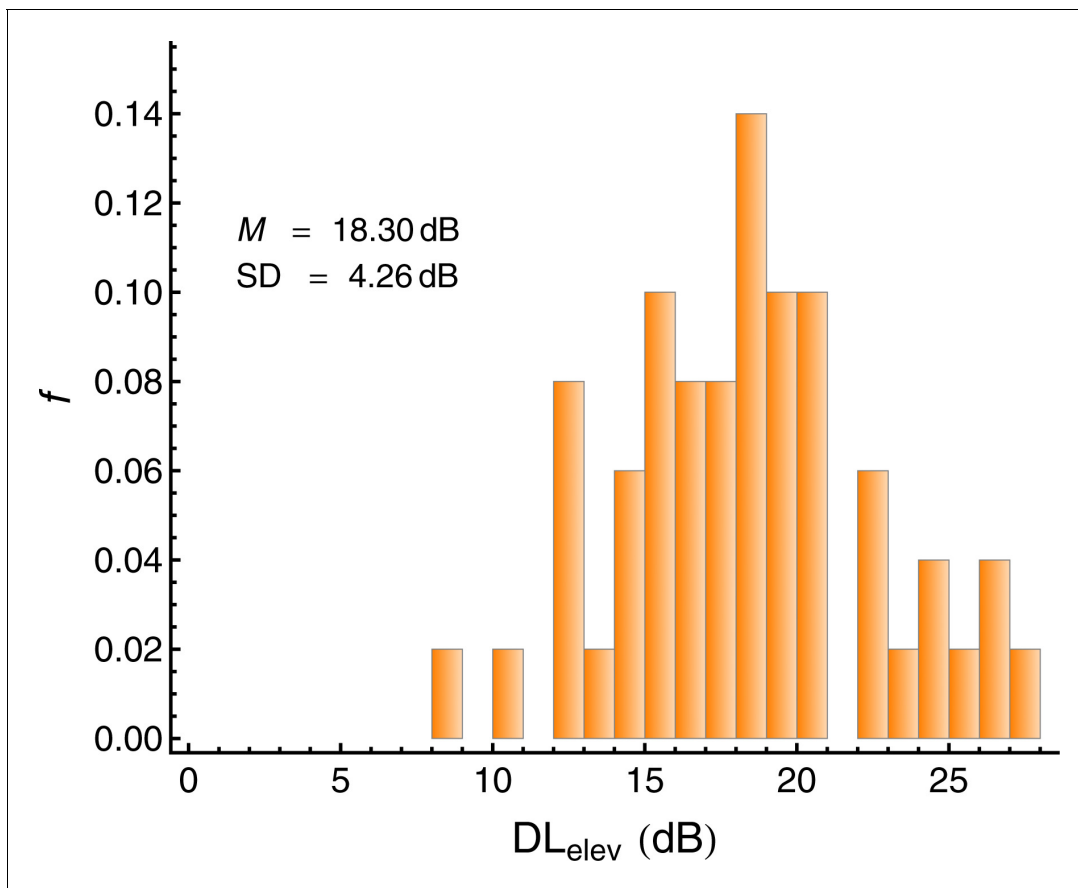


Figure 4. Individual elevation of the intensity difference limen caused by the backward maskers (DL_{elev}), defined as the difference between the DL under masking and the DL in quiet. Lower values represent a better capability of directing auditory selective attention to the target tones. $N = 50$. DOI: 10.7554/eLife.16747.008

was 0° for the first and the third tone. If the listener is sensitive to the change in IPD, then the four tones are perceived as changing in lateralization. In the other interval, all tones were presented with an interaural phase difference of $IPD = 0^\circ$, corresponding to no change in lateralization from tone to tone. All tones were presented at 30 dB SL with a steady-state duration of 300 ms and 50 ms cosine-squared on- and offset ramps, 20 ms pauses between the tones within an interval, and 200 ms silence between the two intervals. The task was to identify the interval which contained the tones with the phase shift and thus elicited the perception of a location change. Visual trial-by-trial feedback was provided. The initial phase shift was $IPD = 180^\circ$ and was divided by $a = (1.25)^2$ in case of three consecutive correct responses, or multiplied by a after an incorrect response (three down, one up rule). After the third reversal, the step size was reduced to $a = 1.25$. The experimental block ended when nine reversals had been collected or 70 trials had been presented. The geometric mean of the IPD at the last six reversals was taken as the IPD threshold. After a brief practice block, two threshold estimates were obtained. Adaptive tracks in which the SD of the \log_{10} -transformed values of the IPD at the counting reversals was higher than 0.3 or where less than 4 reversals had been collected were excluded from the analysis, which affected only 2 of the 100 tracks. The arithmetic mean of the IPD threshold obtained in the two blocks presenting the TFS-LF test was used as a predictor in the regression analyses (TFS_{th}), representing *sensitivity for the temporal fine structure*. The reliability of TFS_{th} across the two measurements (blocks) was $ICC(A,2) = 0.682$. **Figure 5** shows the distribution of TFS_{th} . As expected, there was considerable inter-individual variation of the binaural TFS sensitivity, compatible with previous reports of both monaural (Ruggles et al., 2011) and binaural TFS sensitivity (Füllgrabe, 2013; Ross et al., 2007).

Flanker task (visual selective attention)

To measure spatial visual selective attention, a flanker task as established by *Eriksen and Eriksen (1974)* was used, in a variant proposed by *Kramer and Jacobson (1991)*. The participants' task was to decide whether a target line presented on a computer screen was dotted or dashed. The target line was presented in vertical orientation, on the center of the display. It was surrounded by other lines, the so-called flankers. In one condition (*Figure 6*, right column), the two vertical lines adjacent to the target line were associated with the *incompatible* response. If the target line was dashed, the distractor lines were dotted, and vice versa. In a control condition (*Figure 6*, left column), the adjacent distractor lines were solid, and thus not associated with one of the responses relevant for the target line, this is the *neutral condition*. The two flanker lines adjacent to the target line were either connected with the target line with horizontal solid lines (*Figure 6*, upper row), or they were connected with two additional, vertically oriented solid lines (*Figure 6*, lower row). In the former condition ('same object'), the target line and the distractor lines can be expected to be perceived as belonging to the same visual object (*Kramer and Jacobson, 1991*). In the latter condition ('different object'), the target line and the distractors should be grouped into separate objects. According to the concept of object-based attention (e.g., *Kahneman et al., 1981*), ignoring the flankers should be more difficult if the flankers and the target are perceived as belonging to the same object. To further emphasize the grouping, the target line and the adjacent flankers were presented in the same color in the same-object condition, and in different colors in the different-object condition. The colors blue and green were used, and the target line was equally often presented in blue and in green.

The stimuli were presented on a CRT display (frame rate 85 Hz), with a viewing distance of 100 cm. The size of the vertical lines was 0.9° of visual angle (vertical) by 0.04° (horizontal). The horizontal separation between the lines was 0.25° . The trial started with a blank gray screen presented for 500

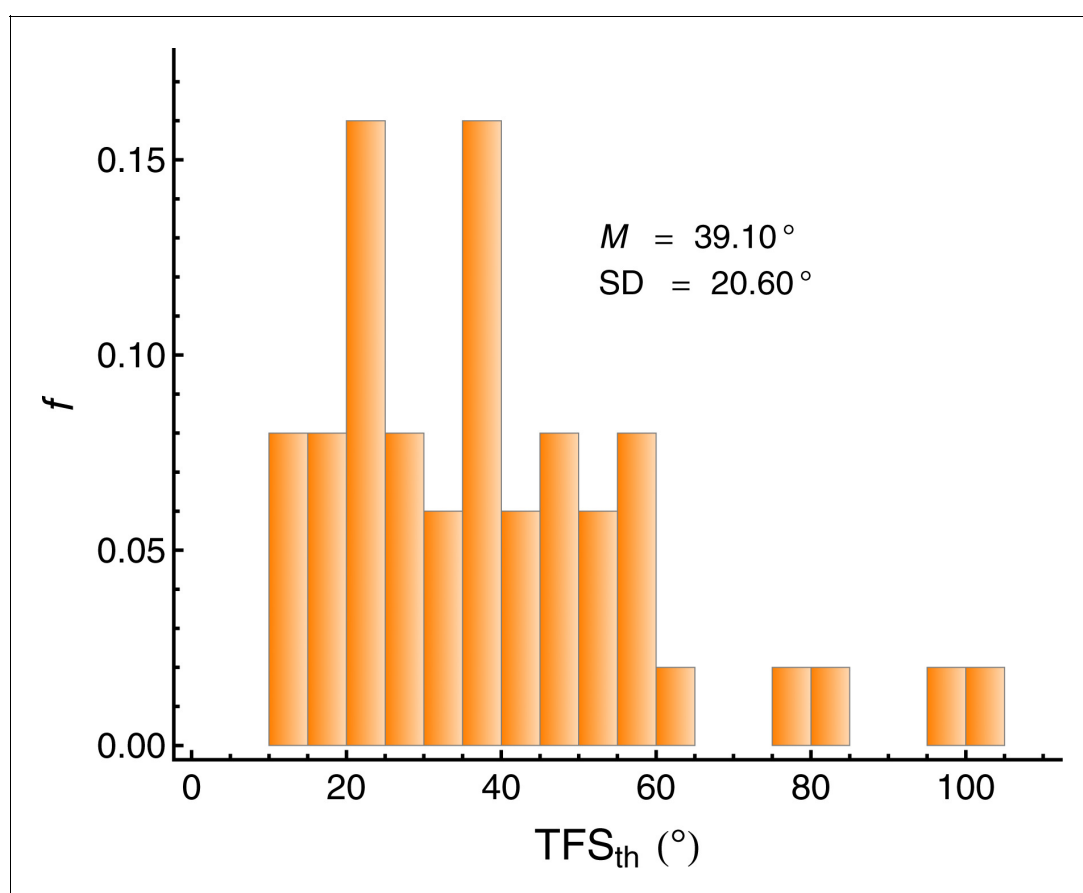


Figure 5. Individual IPD thresholds in the TFS-LF test (TFS_{th}). Lower values represent better binaural sensitivity for the temporal fine structure. $N = 50$. DOI: 10.7554/eLife.16747.009

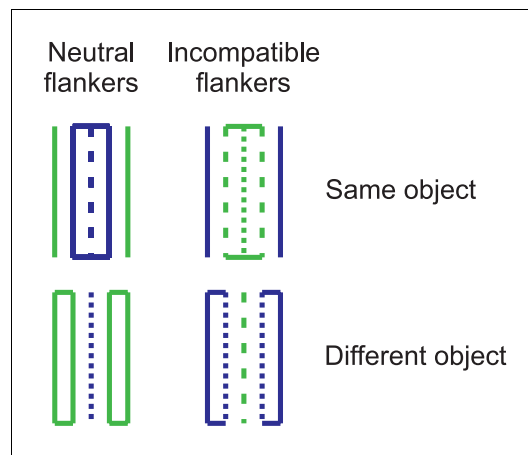


Figure 6. Example stimuli from the flanker task (*Kramer and Jacobson, 1991*) used to measure spatial visual selective attention. The participants' task was to decide whether the central *target line* was dotted or dashed. In the *neutral condition* (left column), the neighboring flanker lines were solid and therefore not associated with a response. In the *incompatible condition* (right column), the flanker lines were associated with the incompatible response. The horizontal lines and the colors promoted the perceptual grouping of the target line and the flankers as either belonging to the *same object* (upper row), or into *different objects* (lower row).

DOI: [10.7554/eLife.16747.010](https://doi.org/10.7554/eLife.16747.010)

ms, followed by a fixation cross presented for 200 ms, after which the stimulus was presented. Participants responded by pressing two designated buttons on a numeric keypad, and received immediate visual feedback concerning the correctness of the response. They first received a practice block containing all of the 16 combinations of target type (dotted, dashed), distractor type (neutral, incompatible), object condition (same object, different object), and target color (blue, green). Then, three experimental blocks were presented. In each block, all of the 16 conditions were presented four times (64 trials/block), in random order.

Response times (RTs) below 200 ms or above 3000 ms were excluded from the analysis, which affected less than 0.1% of the trials. For each subject, the average correct RT on neutral trials (RT_{neutral}) was computed as a measure of *processing speed* (e.g., *Salthouse, 1996*). Because the asymmetric distribution of RTs can cause problems in regression/correlation analyses (e.g., *Bishara and Hittner, 2012*) and repeated-measures ANOVAs (e.g., *Oberfeld and Franke, 2013*), the RTs were log-transformed prior to all analyses. An rmANOVA on the RTs in the neutral condition showed marginally significant effect of block, $F(2, 98) = 2.39, p = 0.097$. The mean RT was significantly higher in the first block than in blocks 2 and 3, representing a practice effect. For this reason, the data from block 1 were excluded from further analyses. The reliability across the two remaining blocks was high, $ICC(A,2) = 0.920$. The average RT on neutral trials in blocks 2 and 3 was used as a predictor (RT_{neutral}).

As a measure of *visual selective attention*, we used the flanker interference, defined as the difference between (log-transformed) correct RTs in the incompatible condition and the neutral condition ($Int_{\text{Flanker}} = RT_{\text{incompatible}} - RT_{\text{neutral}}$), averaged across the same-object and different-object condition and the two blocks (2 and 3). Lower values represent a better capability of directing visual selective attention to the target line (see *Figure 7*). The reliability of Int_{Flanker} across the two measurements (blocks 2 and 3) was lower than desirable, $ICC(A,2) = 0.596$.

Sentence span task (working memory capacity)

Working memory (WM) capacity was measured with a reading span test, originally proposed by *Daneman and Carpenter (1980)*, which is one of the most established working memory span tasks (*Conway et al., 2005*). A computer version was used (*Lewandowsky et al., 2010*). On each trial, the participants saw an alternating sequence of sentences and consonants. The task was to judge the correctness of each sentence and to remember the following consonant for later serial recall. The sentences were taken from the 'easy' variant of the German version (WMC Multilingual,

downloaded from <http://www.psychologie.uzh.ch/fachrichtungen/allgpsy/Software.html>). After a 1.5 s fixation cross, the first semantically correct (e.g., 'Every rabbit has fur.') or incorrect (e.g., 'Tomorrow is in the past.') sentence appeared centrally on the screen. The participants pressed one of two designated buttons on a computer keyboard to classify the sentence as correct or incorrect. On button press, the sentence disappeared and a single consonant was presented centrally for 1 s. After a 100-ms blank interval, the next sentence appeared. Depending on the list length, three to seven of these sentence-consonant sequences were presented. After the complete list had been presented, the participant was asked to type the remembered series of consonants into a response box displayed on the computer screen. The participants were required to type as many letters as were actually presented in the trial. They were informed that the order of letters mattered and were hence instructed to guess if necessary, rather than skip letters that they could not remember. No feedback was provided.

Each participant received two trials for each of the five list lengths (3, 4, 5, 6, and 7), in random order. The proportion of consonants recalled correctly (that is, reproduced in the correct list position), averaged across the 10 lists, was computed for each subject (partial credit scoring; [Conway et al., 2005](#)). The reliability across the two presentations of each list length was acceptable, $ICC(A,2) = 0.759$. The arcsin-sqrt transformed proportion correct on the sentence span task (SS_{pcorr}) was included as a predictor of cocktail-party listening in the regression analyses.

Self-reported hearing problems

Self-reported hearing-related problems in daily life were assessed via the Speech, Spatial and Qualities of Hearing Scale (SSQ) by [Gatehouse and Noble \(2004\)](#); using the German version

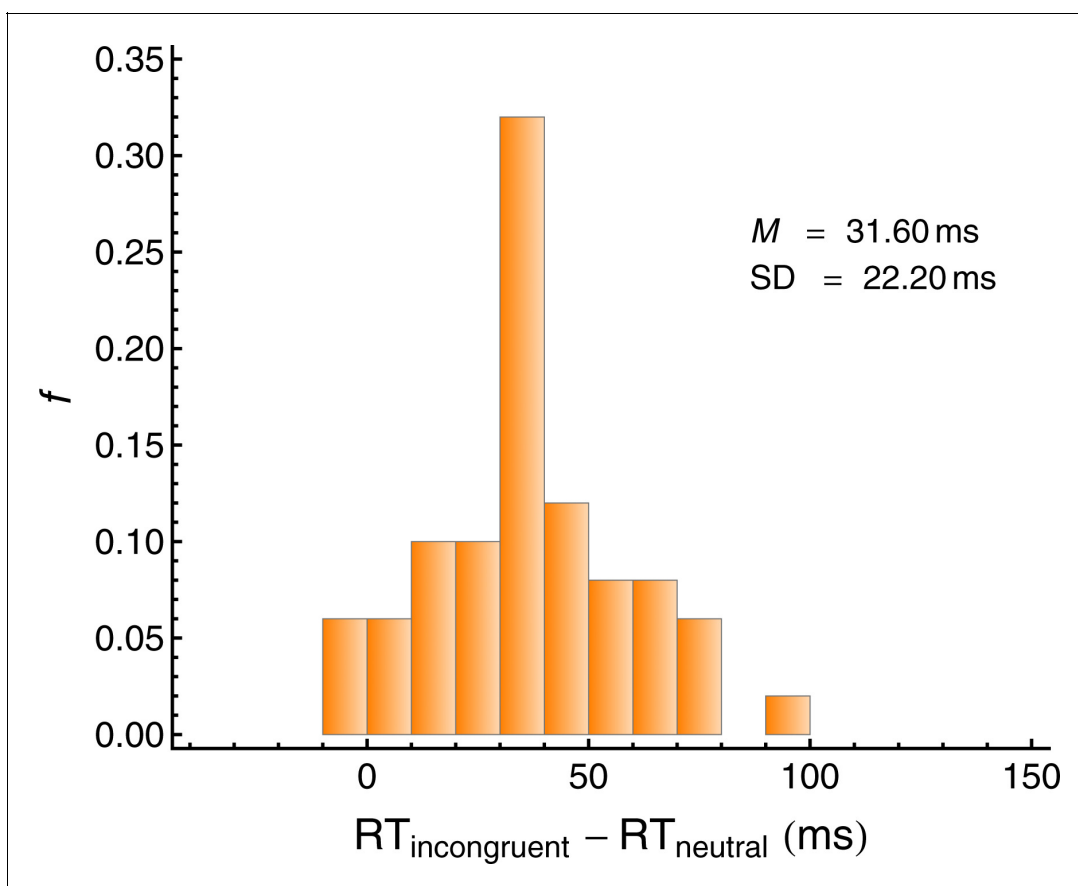


Figure 7. Individual flanker interference ($Int_{Flanker}$) in the visual attention task. Lower values represent a better capability of directing visual selective attention to the target line. $N = 50$.

DOI: [10.7554/eLife.16747.011](https://doi.org/10.7554/eLife.16747.011)

(Kießling *et al.*, 2011). The *Speech hearing* subscale covers speech understanding in the presence of additional speakers (e.g., 'You are in conversation with one person in a room where there are many other people talking. Can you follow what the person you are talking to is saying?'), and is therefore directly relevant for our research question. The *Spatial hearing* subscale indexes the capability of locating static or moving sound sources (e.g., 'You are sitting around a table or at a meeting with several people. You can't see everyone. Can you tell where any person is as soon as they start speaking?'). The *Other qualities* subscale addresses aspects of segregation of sounds, recognition, clarity/naturalness, and listening effort. The SSQ response scales range from 0 to 10, and 10 represents the highest self-rated hearing ability. The mean ratings (with SDs in parentheses) on the Speech hearing, Spatial hearing, and Other qualities scale were 7.16 (1.43), 6.95 (1.39), and 7.82 (1.18), respectively. The SSQ total score, which can range between 0 and 30, varied between 12.5 and 29.4 ($M = 21.9$, $SD = 3.6$). Thus, the participants showed considerable variation in their self-reported hearing abilities.

Procedure

Each participant was tested on all tasks. To minimize inter-individual variation due to different task orders, a fixed sequence of tasks was presented. After informed written consent and basic instructions, the experiment started with the measurement of audiometric thresholds, followed by intensity discrimination in quiet, intensity discrimination under backward masking, the cocktail-party listening task, the TFS-LF test, a questionnaire concerning demographic information, the flanker task, the sentence span task, and the SSQ questionnaire. Each task was preceded by detailed instructions and practice trials. The duration of the experimental session was approximately 3 hr, including several short breaks.

Regression analysis

Multiple linear regression was used to analyze the association between the psychoacoustic and cognitive predictors and the speech recognition score (SRS) in the cocktail-party listening task. As explained above, proportions (SRS and $SS_{P_{corr}}$) were arcsin-sqrt transformed, and the response-time measures ($RT_{neutral}$, $Int_{Flanker}$) were based on log-transformed RTs. All variables were z-standardized. The nine predictors were entered simultaneously. Following the recommendations by *Belsley et al. (1980)*; we analyzed the externally studentized residuals, and the DFFITS index proposed by *Belsley et al. (1980)* as a measure of the influence of an observation. Observations for which the absolute value of the externally studentized residual exceeded 1.96 or with an absolute DFFITS value exceeding $2\sqrt{p/N}$ (where $N = 50$ is the number of subjects, and $p=9$ is the number of predictors) were defined as outliers. This resulted in the exclusion of 5 of the 50 subjects from the regression analysis. The maximum condition index (*Belsley et al., 1980*) was 2.49. *Belsley et al. (1980)* suggested that only condition indices of at least 30 indicate potential problems with multicollinearity. It should be noted that according to the Gauß-Markov theorem (*Gauss, 1821*) the estimates provided by the multiple regression analysis will remain unbiased in the presence of correlated predictors. However, multicollinearity could inflate the variance of the estimated regression coefficients (e.g., *Greene, 2008*), resulting in non-significant regression coefficients.

Q-Q plots of the residuals showed no systematic deviations from normality, and plots of the SRS as a function of the predictors showed no severe deviations from linearity. Thus, linear multiple regression was an appropriate method to assess the influence of the nine predictors on the speech recognition score, and to gauge their relative importance. Note that unlike most previous studies on factors influencing cocktail-party listening, our analyses did not focus on pairwise correlations, because only multiple regression provides valid information about the effects of multiple, partly correlated predictors (see *Table 2*).

Data are available from the Dryad Digital Repository (*Oberfeld and Klöckner-Nowotny, 2016*).

Acknowledgements

This work was supported by a grant from Deutsche Forschungsgemeinschaft (www.dfg.de) to Daniel Oberfeld (OB 346/4-2: "Temporal aspects of auditory intensity processing"). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the

manuscript. The authors have no financial relationship with the organization that sponsored the research. No additional external funding received. The authors declare that they have no conflict of interest. We are grateful to Marius Frenken for his assistance in preparing the figures, and to Annika Grotjohann, Marius Frenken, Julia Pfeiff, and Jannis Renner for their help with data collection. We thank Eve Marder, Barbara Shinn-Cunningham, Hari Bharadwaj and an anonymous reviewer for helpful comments on a previous version of this paper.

Additional information

Funding

| Funder | Grant reference number | Author |
|---------------------------------|------------------------|-----------------|
| Deutsche Forschungsgemeinschaft | OB 346/4-2 | Daniel Oberfeld |

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

DO, FK-N, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article

Author ORCIDs

Daniel Oberfeld,  <http://orcid.org/0000-0002-6710-3309>

Ethics

Human subjects: The experiment was conducted according to the principles expressed in the Declaration of Helsinki. All listeners participated voluntarily after providing informed written consent, after the topic of the study and potential risks had been explained to them. They were uninformed about the experimental hypotheses. The study was approved by the ethics committee of the Department of Psychology, Johannes Gutenberg-Universitaet Mainz.

Additional files

Major datasets

The following dataset was generated:

| Author(s) | Year | Dataset title | Dataset URL | Database, license, and accessibility information |
|--------------------------------|------|--|---|--|
| Oberfeld D, Klöckner-Nowotny F | 2016 | Experimental data on psychoacoustic and cognitive predictors of cocktail-party listening | http://dx.doi.org/10.5061/dryad.f96cr | Available at Dryad Digital Repository under a CC0 Public Domain Dedication |

References

- Akeroyd MA. 2008. Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology* **47**:S53–S71. doi: [10.1080/14992020802301142](https://doi.org/10.1080/14992020802301142), PMID: [19012113](https://pubmed.ncbi.nlm.nih.gov/19012113/)
- Azen R, Budescu DV. 2003. The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods* **8**:129–148. doi: [10.1037/1082-989X.8.2.129](https://doi.org/10.1037/1082-989X.8.2.129), PMID: [12924811](https://pubmed.ncbi.nlm.nih.gov/12924811/)
- Bartlett MS. 1936. The square root transformation in analysis of variance. *Supplement to the Journal of the Royal Statistical Society* **3**:68–78. doi: [10.2307/2983678](https://doi.org/10.2307/2983678)
- Belsley DA, Kuh E, Welsch RA. 1980. Regression diagnostics: Identifying influential data and sources of collinearity. *Wiley Series in Probability and Statistics*. Hoboken, N.J: Wiley. doi: [10.1002/0471725153](https://doi.org/10.1002/0471725153)
- Best V, Ozmeral EJ, Shinn-Cunningham BG. 2007. Visually-guided attention enhances target identification in a complex auditory scene. *Journal of the Association for Research in Otolaryngology* **8**:294–304. doi: [10.1007/s10162-007-0073-z](https://doi.org/10.1007/s10162-007-0073-z), PMID: [17453308](https://pubmed.ncbi.nlm.nih.gov/17453308/)

- Bharadwaj HM**, Masud S, Mehraei G, Verhulst S, Shinn-Cunningham BG. 2015. Individual differences reveal correlates of hidden hearing deficits. *Journal of Neuroscience* **35**:2161–2172. doi: [10.1523/JNEUROSCI.3915-14.2015](https://doi.org/10.1523/JNEUROSCI.3915-14.2015), PMID: [25653371](https://pubmed.ncbi.nlm.nih.gov/25653371/)
- Bishara AJ**, Hittner JB. 2012. Testing the significance of a correlation with nonnormal data: comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological Methods* **17**:399–417. doi: [10.1037/a0028087](https://doi.org/10.1037/a0028087), PMID: [22563845](https://pubmed.ncbi.nlm.nih.gov/22563845/)
- Bowie CR**, Harvey PD. 2006. Administration and interpretation of the Trail Making Test. *Nature Protocols* **1**: 2277–2281. doi: [10.1038/nprot.2006.390](https://doi.org/10.1038/nprot.2006.390), PMID: [17406468](https://pubmed.ncbi.nlm.nih.gov/17406468/)
- Bregman AS**. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Mass: MIT Press.
- Bronkhorst AW**, Plomp R. 1989. Binaural speech intelligibility in noise for hearing-impaired listeners. *The Journal of the Acoustical Society of America* **86**:1374–1383. doi: [10.1121/1.398697](https://doi.org/10.1121/1.398697), PMID: [2808911](https://pubmed.ncbi.nlm.nih.gov/2808911/)
- Bronkhorst AW**. 2000. The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acustica* **86**:117–128.
- Bronkhorst AW**. 2015. The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Attention, Perception, & Psychophysics* **77**:1465–1487. doi: [10.3758/s13414-015-0882-9](https://doi.org/10.3758/s13414-015-0882-9), PMID: [25828463](https://pubmed.ncbi.nlm.nih.gov/25828463/)
- Budescu DV**. 1993. Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin* **114**:542–551. doi: [10.1037/0033-2909.114.3.542](https://doi.org/10.1037/0033-2909.114.3.542)
- Buschman TJ**, Miller EK. 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* **315**:1860–1862. doi: [10.1126/science.1138071](https://doi.org/10.1126/science.1138071), PMID: [17395832](https://pubmed.ncbi.nlm.nih.gov/17395832/)
- Byrne D**, Dillon H, Tran K, Arlinger S, Wilbraham K, Cox R, Hagerman B, Hetu R, Kei J, Lui C, Kiessling J, Kotby MN, Nasser NHA, El Kholly WAH, Nakanishi Y, Oyer H, Powell R, Stephens D, Meredith R, Sirimanna T, et al. 1994. An international comparison of long-term average speech spectra. *The Journal of the Acoustical Society of America* **96**:2108–2120. doi: [10.1121/1.410152](https://doi.org/10.1121/1.410152)
- Békésy Gv**. 1947. A new audiometer. *Acta Oto-Laryngologica* **35**:411–422. doi: [10.3109/00016484709123756](https://doi.org/10.3109/00016484709123756)
- Cahana-Amitay D**, Spiro A, Sayers JT, Oveis AC, Higby E, Ojo EA, Duncan S, Goral M, Hyun J, Albert ML, Obler LK. 2016. How older adults use cognition in sentence-final word recognition. *Aging, Neuropsychology, and Cognition* **23**:418–444. doi: [10.1080/13825585.2015.1111291](https://doi.org/10.1080/13825585.2015.1111291)
- Carlyon RP**. 2004. How the brain separates sounds. *Trends in Cognitive Sciences* **8**:465–471. doi: [10.1016/j.tics.2004.08.008](https://doi.org/10.1016/j.tics.2004.08.008), PMID: [15450511](https://pubmed.ncbi.nlm.nih.gov/15450511/)
- Cherry EC**. 1953. Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America* **25**:975–979. doi: [10.1121/1.1907229](https://doi.org/10.1121/1.1907229)
- Choi I**, Wang L, Bharadwaj H, Shinn-Cunningham B. 2014. Individual differences in attentional modulation of cortical responses correlate with selective attention performance. *Hearing Research* **314**:10–19. doi: [10.1016/j.heares.2014.04.008](https://doi.org/10.1016/j.heares.2014.04.008), PMID: [24821552](https://pubmed.ncbi.nlm.nih.gov/24821552/)
- Cohen J**. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn Hillsdale, N.J: Lawrence Erlbaum Associates . doi: [10.4324/9780203771587](https://doi.org/10.4324/9780203771587)
- Conway AR**, Cowan N, Bunting MF. 2001. The cocktail party phenomenon revisited: the importance of working memory capacity. *Psychonomic Bulletin & Review* **8**:331–335. doi: [10.3758/BF03196169](https://doi.org/10.3758/BF03196169), PMID: [11495122](https://pubmed.ncbi.nlm.nih.gov/11495122/)
- Conway ARA**, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, Engle RW. 2005. Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review* **12**:769–786. doi: [10.3758/BF03196772](https://doi.org/10.3758/BF03196772)
- Corbetta M**, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience* **3**:201–215. doi: [10.1038/nrn755](https://doi.org/10.1038/nrn755), PMID: [11994752](https://pubmed.ncbi.nlm.nih.gov/11994752/)
- Coull JT**, Nobre AC. 1998. Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience* **18**:7426–7435. PMID: [9736662](https://pubmed.ncbi.nlm.nih.gov/9736662/)
- Cowan N**, Elliott EM, Scott Saults J, Morey CC, Mattox S, Hismjatullina A, Conway AR. 2005. On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology* **51**:42–100. doi: [10.1016/j.cogpsych.2004.12.001](https://doi.org/10.1016/j.cogpsych.2004.12.001), PMID: [16039935](https://pubmed.ncbi.nlm.nih.gov/16039935/)
- Culling JF**, Hawley ML, Litovsky RY. 2004. The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. *The Journal of the Acoustical Society of America* **116**:1057–1065. doi: [10.1121/1.1772396](https://doi.org/10.1121/1.1772396), PMID: [15376672](https://pubmed.ncbi.nlm.nih.gov/15376672/)
- Daneman M**, Carpenter PA. 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* **19**:450–466. doi: [10.1016/S0022-5371\(80\)90312-6](https://doi.org/10.1016/S0022-5371(80)90312-6)
- Darwin CJ**, Hukin RW. 1999. Auditory objects of attention: the role of interaural time differences. *Journal of Experimental Psychology: Human Perception and Performance* **25**:617–629. doi: [10.1037/0096-1523.25.3.617](https://doi.org/10.1037/0096-1523.25.3.617), PMID: [10385981](https://pubmed.ncbi.nlm.nih.gov/10385981/)
- David M**, Lavandier M, Grimault N. 2015. Sequential streaming, binaural cues and lateralization. *The Journal of the Acoustical Society of America* **138**:3500–3512. doi: [10.1121/1.4936902](https://doi.org/10.1121/1.4936902), PMID: [26723307](https://pubmed.ncbi.nlm.nih.gov/26723307/)
- Desimone R**, Duncan J. 1995. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience* **18**:193–222. doi: [10.1146/annurev.ne.18.030195.001205](https://doi.org/10.1146/annurev.ne.18.030195.001205), PMID: [7605061](https://pubmed.ncbi.nlm.nih.gov/7605061/)
- Dirks DD**, Bower D. 1970. Effect of forward and backward masking on speech intelligibility. *Journal of the Acoustical Society of America* **47**:1003. doi: [10.1121/1.1911998](https://doi.org/10.1121/1.1911998), PMID: [5443146](https://pubmed.ncbi.nlm.nih.gov/5443146/)
- Eriksen BA**, Eriksen CW. 1974. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics* **16**:143–149. doi: [10.3758/BF03203267](https://doi.org/10.3758/BF03203267)

- Faul F, Erdfelder E, Buchner A, Lang A-G. 2009. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods* **41**:1149–1160. doi: [10.3758/BRM.41.4.1149](https://doi.org/10.3758/BRM.41.4.1149)
- Francis AL. 2010. Improved segregation of simultaneous talkers differentially affects perceptual and cognitive capacity demands for recognizing speech in competing speech. *Attention, Perception, & Psychophysics* **72**: 501–516. doi: [10.3758/APP.72.2.501](https://doi.org/10.3758/APP.72.2.501), PMID: [20139463](https://pubmed.ncbi.nlm.nih.gov/20139463/)
- Füllgrabe C, Moore BC, Stone MA. 2014. Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience* **6**:347. doi: [10.3389/fnagi.2014.00347](https://doi.org/10.3389/fnagi.2014.00347), PMID: [25628563](https://pubmed.ncbi.nlm.nih.gov/25628563/)
- Füllgrabe C, Rosen S. 2016. Investigating the role of working memory in speech-in-noise identification for listeners with normal hearing. In: van Dijk P, Başkent D, Gaudrain E, de Kleine E, Wagner A, Lanting C (eds). *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*. New York: Springer doi: [10.1007/978-3-319-25474-6_4](https://doi.org/10.1007/978-3-319-25474-6_4)
- Füllgrabe C. 2013. Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. *American Journal of Audiology* **22**:313–315. doi: [10.1044/1059-0889\(2013\)12-0070](https://doi.org/10.1044/1059-0889(2013)12-0070), PMID: [23975124](https://pubmed.ncbi.nlm.nih.gov/23975124/)
- Gatehouse S, Akeroyd MA. 2008. The effects of cueing temporal and spatial attention on word recognition in a complex listening task in hearing-impaired listeners. *Trends in Amplification* **12**:145–161. doi: [10.1177/1084713808317395](https://doi.org/10.1177/1084713808317395), PMID: [18567595](https://pubmed.ncbi.nlm.nih.gov/18567595/)
- Gatehouse S, Noble W. 2004. The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology* **43**:85–99. doi: [10.1080/14992020400050014](https://doi.org/10.1080/14992020400050014), PMID: [15035561](https://pubmed.ncbi.nlm.nih.gov/15035561/)
- Gauss CF. 1821. *Theoria combinationis observationum erroribus minimis obnoxiae. Commentationes Societatis Regiae Scientiarum Gottingensis Recentiores*:33–90.
- Greene WH. 2008. *Econometric Analysis*. 6 edn Upper Saddle River, NJ: Pearson.
- Grose JH, Mamo SK. 2010. Processing of temporal fine structure as a function of age. *Ear and Hearing* **31**:755–760. doi: [10.1097/AUD.0b013e3181e627e7](https://doi.org/10.1097/AUD.0b013e3181e627e7), PMID: [20592614](https://pubmed.ncbi.nlm.nih.gov/20592614/)
- Han LA, Poulsen T. 1998. Equivalent threshold sound pressure levels for Sennheiser HDA 200 earphone and Etymotic Research ER-2 insert earphone in the frequency range 125 Hz to 16 kHz. *Scandinavian Audiology* **27**: 105–112. doi: [10.1080/010503998420342](https://doi.org/10.1080/010503998420342), PMID: [9638829](https://pubmed.ncbi.nlm.nih.gov/9638829/)
- Harrell F. 2015. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Springer . doi: [10.1007/978-3-319-19425-7](https://doi.org/10.1007/978-3-319-19425-7)
- Hartmann WM. 2005. *Signals, sound, and sensation. Modern Acoustics And. Signal Processing*. 5th edn New York: Springer.
- Hastie T, Tibshirani R, Friedman J. 2009. *The elements of statistical learning. Data Mining, Inference, and Prediction*. Second Edn Berlin, New York: Springer.
- Heinrich A, Henshaw H, Ferguson MA. 2015. The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. *Frontiers in Psychology* **6**:782. doi: [10.3389/fpsyg.2015.00782](https://doi.org/10.3389/fpsyg.2015.00782), PMID: [26136699](https://pubmed.ncbi.nlm.nih.gov/26136699/)
- Hopkins K, Moore BC. 2010. Development of a fast method for measuring sensitivity to temporal fine structure information at low frequencies. *International Journal of Audiology* **49**:940–946. doi: [10.3109/14992027.2010.512613](https://doi.org/10.3109/14992027.2010.512613), PMID: [20874427](https://pubmed.ncbi.nlm.nih.gov/20874427/)
- IEC 318. 1970. *An IEC Artificial Ear, of the Wide Band Type, for the Calibration of Earphones Used in Audiometry*. Geneva: International Electrotechnical Commission.
- Ihlefeld A, Shinn-Cunningham B. 2008. Disentangling the effects of spatial cues on selection and formation of auditory objects. *Journal of the Acoustical Society of America* **124**:2224–2235. doi: [10.1121/1.2973185](https://doi.org/10.1121/1.2973185), PMID: [19062861](https://pubmed.ncbi.nlm.nih.gov/19062861/)
- James G, Witten G, Hastie G. 2013. *An Introduction to Statistical Learning*. Vol 103 New York: Springer Science +Business Media . doi: [10.1007/978-1-4614-7138-7](https://doi.org/10.1007/978-1-4614-7138-7)
- Jonides J, Yantis S. 1988. Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics* **43**:346–354. doi: [10.3758/BF03208805](https://doi.org/10.3758/BF03208805), PMID: [3362663](https://pubmed.ncbi.nlm.nih.gov/3362663/)
- Kahneman D, Henik A. 1981. Perceptual organization and attention. In: Kubovy M, Pomerantz J. R (eds). *Perceptual Organization*. Hillsdale, NJ: Erlbaum. p 181–211.
- Kayser H, Ewert SD, Anemüller J, Rohdenburg T, Hohmann V, Kollmeier B. 2009. Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. *EURASIP Journal on Advances in Signal Processing* **2009**:298605. doi: [10.1155/2009/298605](https://doi.org/10.1155/2009/298605)
- Kiang NYS, Watanabe T, Thomas EC, Clark LF. 1965. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. Cambridge, Mass: M.I.T. Press.
- Kidd G, Arbogast TL, Mason CR, Gallun FJ. 2005. The advantage of knowing where to listen. *The Journal of the Acoustical Society of America* **118**:3804–3815. doi: [10.1121/1.2109187](https://doi.org/10.1121/1.2109187), PMID: [16419825](https://pubmed.ncbi.nlm.nih.gov/16419825/)
- Kießling J, Grugel L, Meister H, Meis M. 2011. German translations of the questionnaires SADL, ECHO and SSQ and their evaluation. *Zeitschrift Für Audiologie/Audiological Acoustics* **50**:6–16.
- King A, Hopkins K, Plack CJ. 2014. The effects of age and hearing loss on interaural phase difference discrimination. *The Journal of the Acoustical Society of America* **135**:342–351. doi: [10.1121/1.4838995](https://doi.org/10.1121/1.4838995), PMID: [24437774](https://pubmed.ncbi.nlm.nih.gov/24437774/)
- Kitterick PT, Bailey PJ, Summerfield AQ. 2010. Benefits of knowing who, where, and when in multi-talker listening. *The Journal of the Acoustical Society of America* **127**:2498–2508. doi: [10.1121/1.3327507](https://doi.org/10.1121/1.3327507), PMID: [20370032](https://pubmed.ncbi.nlm.nih.gov/20370032/)

- Kramer AF**, Jacobson A. 1991. Perceptual organization and focused attention: the role of objects and proximity in visual processing. *Perception & Psychophysics* **50**:267–284. doi: [10.3758/BF03206750](https://doi.org/10.3758/BF03206750), PMID: [1754368](https://pubmed.ncbi.nlm.nih.gov/1754368/)
- Kujawa SG**, Liberman MC. 2009. Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. *Journal of Neuroscience* **29**:14077–14085. doi: [10.1523/JNEUROSCI.2845-09.2009](https://doi.org/10.1523/JNEUROSCI.2845-09.2009), PMID: [19906956](https://pubmed.ncbi.nlm.nih.gov/19906956/)
- Lavandier M**, Culling JF. 2010. Prediction of binaural speech intelligibility against noise in rooms. *The Journal of the Acoustical Society of America* **127**:387–399. doi: [10.1121/1.3268612](https://doi.org/10.1121/1.3268612), PMID: [20058985](https://pubmed.ncbi.nlm.nih.gov/20058985/)
- Lebreton JM**, Ployhart RE, Ladd RT. 2004. A monte carlo comparison of relative importance methodologies. *Organizational Research Methods* **7**:258–282. doi: [10.1177/1094428104266017](https://doi.org/10.1177/1094428104266017)
- Levitt H**. 1971. Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America* **49**:467–477. doi: [10.1121/1.1912375](https://doi.org/10.1121/1.1912375), PMID: [5541744](https://pubmed.ncbi.nlm.nih.gov/5541744/)
- Lewandowsky S**, Oberauer K, Yang LX, Ecker UK. 2010. A working memory test battery for MATLAB. *Behavior Research Methods* **42**:571–585. doi: [10.3758/BRM.42.2.571](https://doi.org/10.3758/BRM.42.2.571), PMID: [20479189](https://pubmed.ncbi.nlm.nih.gov/20479189/)
- Mattys SL**, Davis MH, Bradlow AR, Scott SK. 2012. Speech recognition in adverse conditions: A review. *Language and Cognitive Processes* **27**:953–978. doi: [10.1080/01690965.2012.705006](https://doi.org/10.1080/01690965.2012.705006)
- McGraw KO**, Wong SP. 1996. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* **1**:30–46. doi: [10.1037/1082-989X.1.1.30](https://doi.org/10.1037/1082-989X.1.1.30)
- Michalka SW**, Kong L, Rosen ML, Shinn-Cunningham BG, Somers DC. 2015. Short-term memory for space and time flexibly recruit complementary sensory-biased frontal lobe attention networks. *Neuron* **87**:882–892. doi: [10.1016/j.neuron.2015.07.028](https://doi.org/10.1016/j.neuron.2015.07.028), PMID: [26291168](https://pubmed.ncbi.nlm.nih.gov/26291168/)
- Moore BC**. 2008. The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology* **9**:399–406. doi: [10.1007/s10162-008-0143-x](https://doi.org/10.1007/s10162-008-0143-x), PMID: [18855069](https://pubmed.ncbi.nlm.nih.gov/18855069/)
- Moray N**. 1959. Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology* **11**:56–60. doi: [10.1080/17470215908416289](https://doi.org/10.1080/17470215908416289)
- Neher T**, Behrens T, Carlile S, Jin C, Kragelund L, Petersen AS, Schaik A. 2009. Benefit from spatial separation of multiple talkers in bilateral hearing-aid users: Effects of hearing loss, age, and cognition. *International Journal of Audiology* **48**:758–774. doi: [10.3109/14992020903079332](https://doi.org/10.3109/14992020903079332), PMID: [19951144](https://pubmed.ncbi.nlm.nih.gov/19951144/)
- Neher T**, Laugesen S, Jensen NS, Kragelund L. 2011. Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *The Journal of the Acoustical Society of America* **130**:1542–1558. doi: [10.1121/1.3608122](https://doi.org/10.1121/1.3608122), PMID: [21895093](https://pubmed.ncbi.nlm.nih.gov/21895093/)
- Neher T**, Lunner T, Hopkins K, Moore BC. 2012. Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners (L). *The Journal of the Acoustical Society of America* **131**:2561–2564. doi: [10.1121/1.3689850](https://doi.org/10.1121/1.3689850), PMID: [22501036](https://pubmed.ncbi.nlm.nih.gov/22501036/)
- Oberfeld D**, Franke T. 2013. Evaluating the robustness of repeated measures analyses: The case of small sample sizes and nonnormal data. *Behavior Research Methods* **45**:792–812. doi: [10.3758/s13428-012-0281-2](https://doi.org/10.3758/s13428-012-0281-2)
- Oberfeld D**, Klöckner-Nowotny F. 2016. Experimental data on psychoacoustic and cognitive predictors of cocktail-party listening. *Dryad Digital Repository*. doi: [10.5061/dryad.f96cr](https://doi.org/10.5061/dryad.f96cr)
- Oberfeld D**, Stahn P, Kuta M. 2012. Binaural release from masking in forward-masked intensity discrimination: evidence for effects of selective attention. *Hearing Research* **294**:1–9. doi: [10.1016/j.heares.2012.09.004](https://doi.org/10.1016/j.heares.2012.09.004), PMID: [23010335](https://pubmed.ncbi.nlm.nih.gov/23010335/)
- Oberfeld D**, Stahn P, Kuta M. 2014. Why do forward maskers affect auditory intensity discrimination? Evidence from "molecular psychophysics". *PLoS One* **9**. doi: [10.1371/journal.pone.0099745](https://doi.org/10.1371/journal.pone.0099745), PMID: [24937050](https://pubmed.ncbi.nlm.nih.gov/24937050/)
- Oberfeld D**, Stahn P. 2012. Sequential grouping modulates the effect of non-simultaneous masking on auditory intensity resolution. *PLoS One* **7**:e48054. doi: [10.1371/journal.pone.0048054](https://doi.org/10.1371/journal.pone.0048054), PMID: [23110174](https://pubmed.ncbi.nlm.nih.gov/23110174/)
- Pashler HE**. 1998. *The Psychology of Attention*. Cambridge, Mass: MIT Press.
- Plack CJ**, Barker D, Prendergast G. 2014. Perceptual consequences of "hidden" hearing loss. *Trends in Hearing* **18**:2331216514550621. doi: [10.1177/2331216514550621](https://doi.org/10.1177/2331216514550621), PMID: [25204468](https://pubmed.ncbi.nlm.nih.gov/25204468/)
- Plack CJ**, Viemeister NF. 1992. Intensity discrimination under backward masking. *The Journal of the Acoustical Society of America* **92**:3097–3101. doi: [10.1121/1.404205](https://doi.org/10.1121/1.404205), PMID: [1474224](https://pubmed.ncbi.nlm.nih.gov/1474224/)
- Posner MI**. 1980. Orienting of attention. *The Quarterly Journal of Experimental Psychology* **32**:3–25. PMID: [7367577](https://pubmed.ncbi.nlm.nih.gov/7367577/)
- Pulvermüller F**, Shtyrov Y. 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. *Progress in Neurobiology* **79**:49–71. doi: [10.1016/j.pneurobio.2006.04.004](https://doi.org/10.1016/j.pneurobio.2006.04.004), PMID: [16814448](https://pubmed.ncbi.nlm.nih.gov/16814448/)
- Robertson IH**, Ward T, Ridgeway V, Nimmo-Smith I. 1996. The structure of normal human attention: The Test of Everyday Attention. *Journal of the International Neuropsychological Society* **2**:525–534. doi: [10.1017/S1355617700001697](https://doi.org/10.1017/S1355617700001697), PMID: [9375156](https://pubmed.ncbi.nlm.nih.gov/9375156/)
- Ross B**, Fujioka T, Tremblay KL, Picton TW. 2007. Aging in binaural hearing begins in mid-life: evidence from cortical auditory-evoked responses to changes in interaural phase. *Journal of Neuroscience* **27**:11172–11178. doi: [10.1523/JNEUROSCI.1813-07.2007](https://doi.org/10.1523/JNEUROSCI.1813-07.2007), PMID: [17942712](https://pubmed.ncbi.nlm.nih.gov/17942712/)
- Ruggles D**, Bharadwaj H, Shinn-Cunningham BG. 2011. Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *PNAS* **108**:15516–15521. doi: [10.1073/pnas.1108912108](https://doi.org/10.1073/pnas.1108912108), PMID: [21844339](https://pubmed.ncbi.nlm.nih.gov/21844339/)
- Ruggles D**, Bharadwaj H, Shinn-Cunningham BG. 2012. Why middle-aged listeners have trouble hearing in everyday settings. *Current Biology* **22**:1417–1422. doi: [10.1016/j.cub.2012.05.025](https://doi.org/10.1016/j.cub.2012.05.025), PMID: [22727697](https://pubmed.ncbi.nlm.nih.gov/22727697/)

- Ruggles D**, Shinn-Cunningham B. 2011. Spatial selective auditory attention in the presence of reverberant energy: individual differences in normal-hearing listeners. *Journal of the Association for Research in Otolaryngology* **12**:395–405. doi: [10.1007/s10162-010-0254-z](https://doi.org/10.1007/s10162-010-0254-z), PMID: [21128091](https://pubmed.ncbi.nlm.nih.gov/21128091/)
- Sach AJ**, Hill NI, Bailey PJ. 2000. Auditory spatial attention using interaural time differences. *Journal of Experimental Psychology: Human Perception and Performance* **26**:717–729. doi: [10.1037/0096-1523.26.2.717](https://doi.org/10.1037/0096-1523.26.2.717), PMID: [10811171](https://pubmed.ncbi.nlm.nih.gov/10811171/)
- Salthouse TA**. 1996. The processing-speed theory of adult age differences in cognition. *Psychological Review* **103**:403–428. doi: [10.1037/0033-295X.103.3.403](https://doi.org/10.1037/0033-295X.103.3.403), PMID: [8759042](https://pubmed.ncbi.nlm.nih.gov/8759042/)
- Salthouse TA**. 2000. Aging and measures of processing speed. *Biological Psychology* **54**:35–54. doi: [10.1016/S0301-0511\(00\)00052-1](https://doi.org/10.1016/S0301-0511(00)00052-1), PMID: [11035219](https://pubmed.ncbi.nlm.nih.gov/11035219/)
- Sander MC**, Lindenberger U, Werkle-Bergner M. 2012. Lifespan age differences in working memory: a two-component framework. *Neuroscience and Biobehavioral Reviews* **36**:2007–2033. doi: [10.1016/j.neubiorev.2012.06.004](https://doi.org/10.1016/j.neubiorev.2012.06.004), PMID: [22771333](https://pubmed.ncbi.nlm.nih.gov/22771333/)
- Schlauch RS**, Lanthier N, Neve J. 1997. Forward-masked intensity discrimination: duration effects and spectral effects. *The Journal of the Acoustical Society of America* **102**:461–467. doi: [10.1121/1.419610](https://doi.org/10.1121/1.419610), PMID: [9228808](https://pubmed.ncbi.nlm.nih.gov/9228808/)
- Schoof T**, Rosen S. 2014. The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Frontiers in Aging Neuroscience* **6**:307. doi: [10.3389/fnagi.2014.00307](https://doi.org/10.3389/fnagi.2014.00307), PMID: [25429266](https://pubmed.ncbi.nlm.nih.gov/25429266/)
- Shinn-Cunningham BG**. 2008. Object-based auditory and visual attention. *Trends in Cognitive Sciences* **12**:182–186. doi: [10.1016/j.tics.2008.02.003](https://doi.org/10.1016/j.tics.2008.02.003), PMID: [18396091](https://pubmed.ncbi.nlm.nih.gov/18396091/)
- Spence CJ**, Driver J. 1994. Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance* **20**:555–574. doi: [10.1037/0096-1523.20.3.555](https://doi.org/10.1037/0096-1523.20.3.555)
- Srinivasan S**, Keil A, Stratis K, Woodruff Carr KL, Smith DW, Carr KLW. 2012. Effects of cross-modal selective attention on the sensory periphery: cochlear sensitivity is altered by selective attention. *Neuroscience* **223**:325–332. doi: [10.1016/j.neuroscience.2012.07.062](https://doi.org/10.1016/j.neuroscience.2012.07.062), PMID: [22871520](https://pubmed.ncbi.nlm.nih.gov/22871520/)
- Stroop JR**. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* **18**:643–662. doi: [10.1037/h0054651](https://doi.org/10.1037/h0054651)
- Styles EA**. 2006. *The Psychology of Attention (2nd ed)*. Hove, UK: Psychology Press
- Thomas DR**, Zumbo BD, Kwan E, Schweitzer L. 2014. On Johnson's (2000) relative weights method for assessing variable importance: A reanalysis. *Multivariate Behavioral Research* **49**:329–338. doi: [10.1080/00273171.2014.905766](https://doi.org/10.1080/00273171.2014.905766), PMID: [26765801](https://pubmed.ncbi.nlm.nih.gov/26765801/)
- Tibshirani R**. 1996. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society Series B-Methodological* **58**:267–288.
- Tibshirani R**. 2011. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B* **73**:273–282. doi: [10.1111/j.1467-9868.2011.00771.x](https://doi.org/10.1111/j.1467-9868.2011.00771.x)
- Tonidandel S**, LeBreton JM. 2011. Relative importance analysis: A useful supplement to regression analysis. *Journal of Business and Psychology* **26**:1–9. doi: [10.1007/s10869-010-9204-3](https://doi.org/10.1007/s10869-010-9204-3)
- Tun PA**, Wingfield A. 1999. One voice too many: Adult age differences in language processing with different types of distracting sounds. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* **54B**:P317–P327. doi: [10.1093/geronb/54B.5.P317](https://doi.org/10.1093/geronb/54B.5.P317)
- van Rooij JC**, Plomp R, Orlebeke JF. 1989. Auditive and cognitive factors in speech perception by elderly listeners. I: Development of test battery. *The Journal of the Acoustical Society of America* **86**:1294–1309. doi: [10.1121/1.398744](https://doi.org/10.1121/1.398744), PMID: [2808905](https://pubmed.ncbi.nlm.nih.gov/2808905/)
- Wagener K**, Kühnel V, Kollmeier B. 1999a. Entwicklung und Evaluation eines Satztests für die deutsche Sprache. I: Design des Oldenburger Satztests. *Zeitschrift Für Audiologie* **38**:4–15.
- Wagener K**, Kühnel V, Kollmeier B. 1999b. Entwicklung und Evaluation eines Satztests für die deutsche Sprache. II: Optimierung des Oldenburger Satztests. *Zeitschrift Für Audiologie* **38**:44–56.
- Wagener K**, Kühnel V, Kollmeier B. 1999c. Entwicklung und Evaluation eines Satztests für die deutsche Sprache, III: Evaluation des Oldenburger Satztests. *Zeitschrift Für Audiologie* **38**:86–95.
- Xiang J**, Simon J, Elhilali M. 2010. Competing streams at the cocktail party: exploring the mechanisms of attention and temporal integration. *Journal of Neuroscience* **30**:12084–12093. doi: [10.1523/JNEUROSCI.0827-10.2010](https://doi.org/10.1523/JNEUROSCI.0827-10.2010), PMID: [20826671](https://pubmed.ncbi.nlm.nih.gov/20826671/)
- Yantis S**, Jonides J. 1990. Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance* **16**:121–134. doi: [10.1037/0096-1523.16.1.121](https://doi.org/10.1037/0096-1523.16.1.121), PMID: [2137514](https://pubmed.ncbi.nlm.nih.gov/2137514/)
- Yost WA**, Dye RH, Sheft S. 1996. A simulated "cocktail party" with up to three sound sources. *Perception & Psychophysics* **58**:1026–1036. doi: [10.3758/BF03206830](https://doi.org/10.3758/BF03206830), PMID: [8920839](https://pubmed.ncbi.nlm.nih.gov/8920839/)
- Zanto TP**, Gazzaley A. 2014. Attention and ageing. In: Nobre A. C, Kastner S (Eds). *The Oxford Handbook of Attention*. Oxford, UK: Oxford University Press. p. 927–971.
- Zekveld AA**, Rudner M, Johnsrude IS, Rönnberg J. 2013. The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *The Journal of the Acoustical Society of America* **134**:2225–2234. doi: [10.1121/1.4817926](https://doi.org/10.1121/1.4817926), PMID: [23967952](https://pubmed.ncbi.nlm.nih.gov/23967952/)
- Zhao F**, Stephens D. 2007. A critical review of King-Kopetzky syndrome: Hearing difficulties, but normal hearing? *Audiological Medicine* **5**:119–124. doi: [10.1080/16513860701296421](https://doi.org/10.1080/16513860701296421)

- Zion Golumbic EM**, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE. 2013. Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron* **77**:980–991. doi: [10.1016/j.neuron.2012.12.037](https://doi.org/10.1016/j.neuron.2012.12.037), PMID: [23473326](https://pubmed.ncbi.nlm.nih.gov/23473326/)
- Zion Golumbic EM**, Poeppel D, Schroeder CE. 2012. Temporal context in speech processing and attentional stream selection: a behavioral and neural perspective. *Brain and Language* **122**:151–161. doi: [10.1016/j.bandl.2011.12.010](https://doi.org/10.1016/j.bandl.2011.12.010), PMID: [22285024](https://pubmed.ncbi.nlm.nih.gov/22285024/)