



Subjective evaluation of communicative effort for younger and older adults in interactive tasks with energetic and informational masking

Valerie Hazan, Outi Tuomainen, Linda Taschenberger

Department of Speech Hearing and Phonetic Sciences, UCL, UK

v.hazan@ucl.ac.uk, o.tuomainen@ucl.ac.uk, l.taschenberger@ucl.ac.uk

Abstract

The impact of energetic (EM) and informational masking (IM) on speech communication is typically evaluated using perception tests that do not involve actual communication. Here, ratings of effort, concentration and degree of interference were obtained for 51 young, middle-aged and older adults after they had completed communicative tasks (Diapix) with another participant in conditions in which no noise, speech-shaped noise, or three voices were heard in the background. They also completed background sensory and cognitive tests and a quality of hearing questionnaire. The EM condition was perceived as less effortful, requiring less concentration and easier to ignore than those involving IM. Effects were generally greater for talkers taking the lead in the interaction. Even though the two older groups were more affected by IM than young adults in a speech in noise perception test, age did not impact on ratings of effort and ability to ignore the noise in the diapix communicative task. Only for concentration ratings, did the Older Adult group give similar ratings in quiet as when EM was present. Together, these results suggest that evaluations that purely assess receptive speech in older adults do not fully represent the impact of sources of interference on speech communication.

Index Terms: speech communication, speech in noise, speaker-listener interactions, listener effort.

1. Introduction

Communication in noisy environments can be difficult and effortful for many individuals and can be particularly so for specific populations such as non-native speakers and older adults. For example, a study involving nearly half a million Biobank participants showed speech in noise thresholds deteriorating in adults aged 50 and above [1]. It has also been well established that the type of noise can differentially affect populations of listeners, with background noise that includes meaningful speech (informational masking IM) causing greater interference than background noise that is purely energetic (energetic masking EM) [2] with this effect exacerbated in children, for example [3]. Not only is speech intelligibility affected when listening in noise but the amount of listening effort is also increased [e.g. 2] due partially to a higher cognitive load, as suggested in the Framework for Understanding Effortful Listening [FUEL, 4].

Two quite different approaches are taken to the assessment of speech communication in noise. Both focus on the ability to perceive speech rather than a broader aspect of communication. The first approach is to evaluate speech communication in noise using standardized tests that either establish a speech reception threshold (SRT) or intelligibility levels for materials such as digits, single words (as in [1]) or sentence materials such as

BKB-SIN [5]. A quite different approach is to ask individuals to rate their perceived degree of communication difficulty in everyday life using a standardised questionnaire such as the Speech, Spatial and Qualities of Hearing Scale [SSQ, 6]. A study investigating correlations between subjective ratings such as collected via SSQ and objective evaluations of speech perception in noise for a cohort of adults with cochlear implants showed these correlations to be weak to moderate [7]. There are some limitations to both these approaches. Standardised tests only evaluate intelligibility for carefully-controlled read-aloud materials and do not involve any form of communication or the variability expected in spontaneous speech. Questionnaires ask participants to report experiences post-hoc and the questions are, by necessity, quite general, so not sensitive to different levels or types of noise.

More recent paradigms have investigated the impact of challenging listening conditions using tasks that involve real communication and exchange of information between pairs of participants. For example, Diapix [8, 9], a ‘spot the difference’ picture task, has been used to investigate communication in challenging conditions in young [10] and older adults [11] and in children aged 9 to 14 [12], focusing on the clear speech adaptations made in various conditions. Another recent approach is the tangram puzzle task developed by Beechey and colleagues [13]. This task, involving pairs of participants, is run in realistic acoustic environments (e.g. library, office, foodcourt). In [14], changes in perceived effort were related to the intensity of the acoustic environment in the background; greater changes in speech production occurred in the presence of realistic environments with higher sound pressure levels (SPL). The focus in that study was on the loudness of the realistic acoustic environments rather than their informational/energetic profile.

In our study, we investigated age effects for perceived effort, concentration and background interference in the presence of energetic and informational masking using DiapixUK materials [9]. We obtained separate ratings for the participant taking the lead in the interactions and for the more passive participant. These ratings were compared to background measures collected for the same participants.

2. Method

2.1. Participants

51 monolingual native speakers of Standard Southern English participated in the study. They were aged between 19-26 years (Younger Adults, YA, N=20, 10 F, Mean age 21.75 years), between 50 and 64 (Older Middle Aged adults, OMA, N=13, 10 F, Mean age 60.5 years) and between 65-76 years (Older Adults, OA, N=18, 10 F, Mean age 70.4 years). For Diapix, participants were tested in pairs from within a same age group

(one participant was excluded due to failing hearing threshold criteria). All participants were tested for normal hearing thresholds (<20 dB HL) across the 0.25-4 kHz range and reported no history of speech and language impairments or neurological trauma. All participants aged over 65 passed the Montreal Cognitive Assessment (MoCA) –screening test [15].

2.2. Materials

2.2.1. Background tests

CCRM test: Participants completed a test to measure their signal-to-noise ratio (SNR) threshold for simple items presented in noise (e.g. ‘show the dog where the red 6 is’). The adaptive CCRM test (described in [16] as ‘WiNiCS test’ and modelled on the Coordinate Response measure [17]), was used, with either a 3 male-talker babble masker (CCRM-BABB) or speech-shaped noise masker (CCRM-SPSN), and a female voice as target. A three-up one-down adaptive procedure was used to vary SNR, tracking the threshold for a 79.4% correct level from the mean of the reversals excluding the first two. The level of the output was fixed at 65 dB SPL. The test ended after eight reversals or after 30 trials.

BKB sentence test: BKB [18] sentences produced by a male speaker were presented in 8-talker babble in an adaptive test tracking the SNR threshold for a 67% intelligibility level. Participants repeated the sentence heard and keywords were scored by the experimenter. The test ended after 25 trials.

Expressive vocabulary: The ‘Word definitions’ subset of the WISC [19] test evaluated verbal knowledge and concept formation. Participants had to provide a definition (e.g. ‘what does migrate mean’) for a series of words. Scores (0-2) given per item were based on sophistication of the definition.

Letter-number sequencing: This subtest from the WAIS [20] evaluated working memory. Participants were given a set of numbers and letters and had to repeat numbers first, in order, starting with the lowest number, then letters in alphabetical order, increasing in difficulty.

SSQ questionnaire: The ‘Speech Spatial Qualities’ standardised questionnaire [6] was used to get an evaluation of speech and hearing in different situations. It includes 49 questions, with a 10 point Likert response scale. Ratings relating to perceived effort in following conversation (Q3.15 in questionnaire), concentration (Q3.14) and ease of ignoring noise (Q3.18, labelled as ‘interference’) were extracted as most relevant to the Diapix ratings described below.

2.2.2. Diapix task

In order to assess individuals’ ability to communicate meaningfully in challenging listening conditions, participant pairs carried out the interactive ‘spot-the-difference’ Diapix task using the DiapixUK picture sets [9] on a desktop PC. They were seated in separate acoustically-shielded rooms, without sight of each other, and communicated via headsets fitted with a cardioid microphone (Beyerdynamic DT297). Participants saw different versions of the same picture scene and told that they had 10 minutes to find the 12 differences between the pictures. To model multitasking in communication, they also completed a secondary task: pushing a button when hearing the sound of a dog barking and inhibit pressing a button when hearing the sound of a bell. One of the talkers (designated ‘Talker A’) was told to lead the interactions. The other (‘Talker

B’) was a more passive participant who mainly responded to queries by Talker A. All participants carried out both roles.

Diapix was carried out in four listening conditions affecting both participants: i) quiet (QUIET), ii) energetic masking with no informational content (speech-shaped noise, SPSN), iii) informational masking that is semantically related to the picture description task (IMRE; i.e., 3 voices talking about the same picture), and iv) informational masking that is semantically unrelated to the task (IMUR; i.e., 3 voices talking about a different picture). The picture and noise condition orders were randomised. Both IMRE and IMUR were 3-talker maskers consisting of a male, a female and a child. Our expectation was that semantically-related content would cause a greater degree of interference. In order to give a more natural listening environment, we used Spatial Audio Simulation System software (Audio 3D) [21] that mimics real room acoustics combined with head-related transfer functions in real-time. The maskers and the voice of the interlocutor were spatially separated by 1 meter from both each other and the “live” talker [22]. The intensity of all three maskers (SPSN, IMUR, IMRE) was normalised to 72 dB SPL and presented over headphones. The intensity level of the speakers was set to approximate 0 dB SNR when speaking normally.

2.2.3. Rating questionnaire

After completing each Diapix task, both participants completed a paper-based questionnaire, answering four questions using a 10 point Likert scale. The questions were:

1. On average, how noisy did you experience the background noise you heard during the task?
2. Could you easily ignore the background noise?
3. Did you have to concentrate very hard to understand your partner?
4. Did you have to put in a lot of effort to understand your partner?

Descriptors were given at each endpoint of the scale and they could place a X at any point along the scale (including gradations between the numbered markers). The scale was ordered in the same way as in the SSQ questionnaire.

Ratings for participants in Talker A role primarily reflect speaker effort while ratings in Talker B role primarily reflect listener effort. Questions 1 and 2 were not asked in QUIET.

3. Results

Data were analysed using repeated-measures ANOVAs with within-subject effects of condition and between-subject effects of age group.

3.1. Background tests

CCRM test: The effect of masker type on SNR thresholds was significant [$F(1,48)=176.78$; $p<.0001$] with better SNR thresholds (i.e. greater tolerance of noise) for the SPSN masker ($M=-6.14$ dB) than the BABB masker ($M=-1.94$). The masker type by age group interaction was significant [$F(2,48)=9.11$; $p<.0001$] but the main effect of age group was not ($p=.08$). Post-hoc t-tests showed that the effect of age group was not significant for the SPSN masker, but that for the BABB masker, better thresholds were obtained for YA ($M=-3.53$) than for OMA ($M=-0.81$) and OA ($M=-1.48$) groups who did not differ. It was therefore the case that IM caused greater interference to

the OMA and OA groups than to YA. Recall that all participants had average pure tone thresholds better than 20 dB HL.

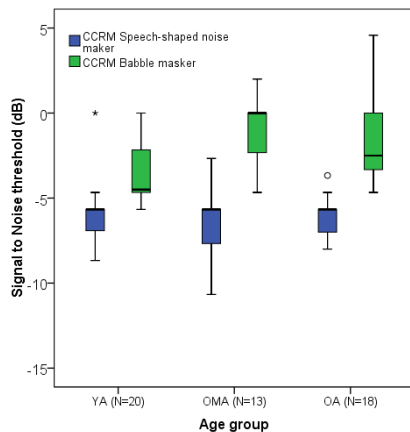


Figure 1: SNR thresholds for CCRM test

BKB sentences: a univariate ANOVA showed that the effect of age group was not significant [$F(2,48)=1.75$; $p=.185$].

Expressive vocabulary: The effect of age group was significant [$F(2,48)=4.07$; $p=.023$]. YA had poorer expressive vocabulary ($M=13.55$) than OMA ($M=15.08$) and OA groups ($M=15.28$) which did not differ.

Letter-number sequencing: YA achieved a higher scaled score ($M=11.7$), revealing better working memory, than the OA group ($M=9.61$) [$F(2,48)=3.45$; $p=.04$]. OMA ($M=10.23$) and OA group scores did not differ.

SSQ questionnaire: The effect of age group was not significant for the ratings of any of the extracted questions.

In summary, in a perception test (CCRM), adults aged 53 and above were more affected by informational masking than young adults. OA participants scored more poorly on letter-number sequencing, reflecting working memory but both OMA and OA groups had better expressive vocabulary than YAs.

3.2. Diapix transaction efficiency

We first examined whether transaction efficiency (i.e., the time taken to find the first eight differences in the picture), varied with age group and condition. The effect of age group just failed to reach significance [$F(2, 48)=3.115$; $p=.053$], with OA ($M=294$ seconds) and OMA ($M=283$) tending to take longer than YA ($M=249$). However, the IM/EM conditions did not take significantly longer to complete than QUIET, suggesting that the acoustic ‘clear speech’ adaptations made by the participants were successful in enabling participants to complete the task successfully.

3.3. Diapix ratings

3.3.1. Effort ratings

Ratings for perceived effort (Question 4) are summarised in Table 1. The effect of condition was significant [$F(3, 144)=69.14$; $p<.0001$]: The QUIET condition was considered less effortful ($M=8.70$) than all conditions involving masking. The semantically-related IM condition (IMRE) ($M=5.22$) and EM (SPSN, $M=5.80$) were rated as similarly effortful and the condition involving semantically-unrelated IM (IMUR) most

effortful ($M=5.06$). There was an effect of Talker role [$F(1, 48)=5.64$; $p=.02$]: participants rated the task as more effortful when in Talker A role ($M=5.95$) than when in passive Talker B role ($M=6.44$). The talker role by age group interaction was significant [$F(2, 48)=3.39$; $p=.04$]: the OMA group reported Talker B role as more effortful which was a reversal of the role ratings for the YA and OA groups. The effect of age group or interaction with condition were not significant suggesting that the presence of EM or IM did not lead older adults to rate the interaction as more effortful than did younger adults.

Table 1: Mean effort ratings (0: Lots of effort – 10: no effort) as a function of condition and talker role (standard deviations in parentheses)

TALKER A ROLE				
Group	QUIET	SPSN	IMRE	IMUR
YA	8.8 (1.8)	4.9 (2.4)	3.8 (1.9)	3.8 (2.1)
OMA	8.8 (1.9)	6.3 (1.8)	5.7 (3.3)	5.1 (3.1)
OA	7.8 (2.7)	6.1 (2.3)	5.2 (2.3)	5.1 (2.5)

TALKER B ROLE				
Group	QUIET	SPSN	IMRE	IMUR
YA	8.8 (1.3)	6.0 (2.2)	5.2 (2.7)	5.6 (2.7)
OMA	9.3 (1.0)	5.3 (3.2)	5.2 (3.0)	5.0 (2.9)
OA	8.7 (1.9)	6.3 (2.5)	6.2 (2.4)	5.8 (2.6)

3.3.2. Concentration ratings

See Table 2 for Diapix ratings for perceived concentration (Q. 3). The effect of condition was significant [$F(3, 144)=58.97$; $p<.0001$]: QUIET was rated as requiring less concentration ($M=8.24$) than all conditions with masking, and IM conditions required more concentration than the EM condition. There was a significant interaction between talker role and age group [$F(2, 48)=4.78$; $p=.013$]: Talker A role required greater concentration than Talker B role for YA and OA, but a reverse effect was seen for OMA group (see Table 2). Finally a condition by age interaction [$F(6, 144)=2.63$; $p=.019$]: post-hocs revealed that for YA and OMA groups, QUIET required less concentration than all masking conditions which did not differ whereas for OA, QUIET was perceived as requiring less concentration than the IM conditions but did not differ from the EM condition.

3.3.3. Interference ratings

The only significant effect was for condition [$F(2, 100)=15.10$; $p<.0001$]: the SPSN condition ($M=6.00$) was rated as being easier to ignore than the two IM conditions (IMRE: 4.87; IMUR: 4.95). The effect of talker role was also significant [$F(1.50)=4.46$; $p=.04$]: the interference was harder to ignore in Talker A role ($M=5.00$) than in Talker B role ($M=5.55$).

In summary, there were clear differences in the effects of EM and IM on perceived difficulty in communicative interactions. The EM condition was perceived as less effortful and requiring less concentration than the semantically-unrelated IM condition. It was also easier to ignore EM than IM. Talkers taking the lead in the interaction reported greater effort, concentration and interference effects than talkers in the more passive listening role. Within this population with normal hearing thresholds, age did not impact on ratings of effort and ability to ignore the noise. Only for concentration ratings, was

the OA group giving a rating that suggested they required as much concentration for communication in quiet as when in the presence of EM.

Table 2: Mean concentration ratings (0: concentrate hard – 10: No need to concentrate) as a function of condition and talker role (standard deviations in parentheses)

TALKER A ROLE

Group	QUIET	SPSN	IMRE	IMUR
YA	8.6 (2.1)	4.4 (2.2)	3.7 (2.5)	3.7 (2.1)
OMA	8.8 (2.2)	5.8 (3.4)	5.6 (3.3)	4.7 (3.1)
OA	7.1 (2.8)	5.7 (2.3)	4.8 (1.8)	4.1 (2.5)

TALKER B ROLE

Group	QUIET	SPSN	IMRE	IMUR
YA	8.6 (1.8)	6.0 (2.1)	4.8 (2.6)	5.3 (2.7)
OMA	8.8 (1.4)	4.8 (3.1)	4.1 (2.8)	4.6 (3.3)
OA	7.4 (2.6)	6.2 (2.4)	5.7 (2.5)	5.4 (2.7)

3.4. Correlation between Diapix ratings and SSQ questionnaire

We examined whether ratings given immediately after completing Diapix were correlated with ratings of communication difficulty as collected via SSQ questionnaire. To reduce the number of comparisons, we calculated a mean for IM ratings averaged over IMRE and IMUR. As there was no age effect for SSQ ratings, correlations were carried out over all participants, with $p \leq .01$ was used as significance level.

For Effort ratings, no correlations were significant for any condition (QUIET, IM, EM), although there was a trend for SSQ effort ratings (Q3.15) to be correlated with diapix IM effort ratings as Talker A ($r=.315$, $n=51$, $p=.024$) and as Talker B ($r=.298$, $n=51$, $p=.033$). For Concentration ratings, SSQ ratings (Q3.14) were correlated with Diapix QUIET concentration ratings for Talker B only ($r=.469$, $n=51$, $p=.001$). For Interference ratings, SSQ ratings (Q3.18) were correlated with Diapix IM interference ratings in Talker B role only ($r=.388$, $n=51$, $p=.005$). Overall, even where significant, correlations between Diapix and SSQ ratings were fairly weak.

3.5. Correlation between diapix ratings and background measures

We carried an exploratory analysis to investigate whether perceptual measures for speech in noise (CCRM, BKB), expressive vocabulary (WISC) and working memory (LN sequencing) could predict Diapix communication difficulty ratings. As ratings of effort, concentration and interference were correlated, we calculated, per participant, two composite measures: ‘IM difficulty’ (averaged over talker roles, IMUR/IMRE and the three rating judgments) and ‘EM difficulty’ (averaged over talker roles and rating judgments). Repeated-measures ANOVAs confirmed that composite IM ratings ($M=5.72$) were significantly less favourable than EM ratings ($M=4.90$) [$F(1, 48)=16.64$; $p<.0001$] but that ratings did not differ across age groups ($p=.689$).

As some independent variables showed age effects, we carried out separate linear stepwise regressions for each age group. For the YA and OMA groups, no predictors were obtained for either

IM or EM difficulty ratings. For OA, for IM difficulty ratings, the final model including WISC and CCRM SPSN accounted for 50.3% of the variance (with WISC accounting for 36.7%). For EM difficulty, the model including BKB BABB and WISC accounted for 41.3% (with BKB BABB accounting for 27.4%). For both IM and EM ratings, better expressive vocabulary and less favourable SNR thresholds in perception tests predicted greater perceived difficulty.

4. Discussion

We examined age effects on the impact of energetic and informational masking on speech communication, using subjective ratings of effort, concentration and interference for a communicative task with controlled masking conditions, together with other background measures.

The CCRM perception test confirmed previous findings that IM causes greater perceptual interference than EM, especially for older adults. In the communicative Diapix task, conditions involving IM were rated as requiring more concentration and causing more interference than the EM condition. Contrary to predictions, the semantically-unrelated IM was rated as more effortful than the semantically-related condition: rather than cause greater-interference, the presence of relevant keywords in IMRE possibly aided participants complete the task. In [13], subjective difficulty ratings in a communicative task were sensitive to the intensity level of the background. Here, for maskers presented at a consistent intensity level, difficulty ratings were dependent on whether there was an informational component in the masking. Older adults did not give higher ratings of perceived effort and background interference than young adults, with only a minor age effect on perceived concentration. In our previous Diapix study with older adults [11], OAs with normal hearing patterned with young adults in terms of their clear speech adaptations in challenging conditions while those with mild presbycusis exhibited greater vocal effort so these results are consistent with that study. The fact that Diapix ratings only showed weak correlation with the SSQ questionnaire may reflect the fact that the SSQ is aimed at respondents with hearing loss.

It is also noteworthy that for the YA and OA groups, being in the ‘lead role’ in the communicative interactions led to greater perceived effort and concentration than in the passive Talker B role. This suggests that standard evaluations that purely assess receptive speech likely underestimate the impact of challenging conditions on speech communication, which involves both speech reception and production, as also suggested in [14]. Speaker effort may often be greater than listener effort, even though the latter is the focus of many studies and models of communication in older adults [e.g., 4].

An intriguing finding when evaluating predictors of Diapix ratings was that, for the OA group only, better expressive vocabulary seemed related to communication being rated as more effortful, in IM maskers especially. This may be due to meaningful speech in the background having a greater distracting effect or imposing a greater cognitive load due to a greater lexical search for such participants, but this conclusion may be premature given that these ratings were obtained for 18 older adults only. Further replication of this effect is required.

5. Acknowledgements

This work was supported by the Economic and Social Research Council [grant number ES/P002803/1].

6. References

- [1] D.R. Moore, M. Edmondson-Jones, P. Dawes, H. Fortnum, A. McCormack, R.H. Pierzycki et al. "Relation between Speech-in-Noise Threshold, Hearing Loss and Cognition from 40–69 Years of Age," *PLoS ONE*, vol. 9, no. 9: e107720, 2014.
- [2] M. Rudner, T. Lunner, T. Behrens, E.S. Thorén, J. Rönnerberg "Working memory capacity may influence perceived effort during aided speech recognition in noise," *Journal of the American Academy of Audiology*, vol. 23, no. 8, pp. 577–589, 2012.
- [3] F.L. Wightman, D.J. Kistler. "Informational masking of speech in children: effects of ipsilateral and contralateral distracters," *J. Acoust. Soc. Am.*, vol. 118, pp. 3164–3176, 2005.
- [4] M.K. Pichora-Fuller, S.E. Kramer, M.A. Eckert, B. Edwards, B.W. Hornsby, L.E. Humes, L. E., et al. "Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL)," *Ear and Hearing*, vol. 37, pp. 5S–27S, 2016.
- [5] Etymotic Research. BKB-SIN speech in noise test version 1.03 (compact disc). Elk Grove Village, IL: Etymotic Research, 2005.
- [6] S. Gatehouse and W. Noble. "The Speech, Spatial and Qualities of Hearing Scale (SSQ)," *International Journal of Audiology*, vol. 43, pp. 85–99, 2004.
- [7] G.G. Ramakers, Y.E. Smulders, A. van Zon, G.A. van Zanten, W. Grolman, and I. Stegeman. "Correlation between subjective and objective hearing tests after unilateral and bilateral cochlear implantation," *BMC Ear Nose Throat Disord.* 17:10, 2017.
- [8] K.J. Van Engen, M. Baese-Berk, R.E. Baker, A. Choi, M. Kim and A.R. Bradlow. "The Wildcat Corpus of Native-and Foreign-accented English: Communicative Efficiency across Conversational Dyads with Varying Language Alignment Profiles," *Language and Speech*, vol. 53(4), pp. 510–540, 2010.
- [9] R. Baker and V. Hazan. "DiapixUK: a task for the elicitation of spontaneous speech dialogs," *Behavior Research Methods*, vol. 43, no. 3, pp. 761–770, 2011.
- [10] V. Hazan, and R. Baker. "Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions," *Journal of the Acoustical Society of America*, vol. 130, no. 4, pp. 2139–2152, 2011.
- [11] V. Hazan, O. Tuomainen, J. Kim, C. Davis, B. Sheffield, D. Brungart "Clear speech adaptations in spontaneous speech produced by young and older adults," *Journal of the Acoustical Society of America*, vol. 144, pp. 1331–1346, 2018.
- [12] V. Hazan, O. Tuomainen, and M. Pettinato "Suprasegmental characteristics of spontaneous speech produced in good and challenging communicative conditions by talkers aged 9 to 14 years old," *Journal of Speech, Hearing and Language Research*, vol. 59, pp. S1596–S1607, 2016.
- [13] T. Beechey, J. M. Buchholz and G. Keidser "Eliciting naturalistic conversations: a method for assessing communication ability, subjective experience, and the impacts of noise and hearing impairment", *Journal of Speech, Hearing and Language Research*, vol. 62, pp. 470–482, 2019.
- [14] T. Beechey, J.M. Buchholz and G. Keidser, G. "Measuring communication difficulty through effortful speech production during conversation," *Speech Communication*, vol. 100, pp. 18–29, 2018.
- [15] Z.S. Nasreddine, N.A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J.L. Cummings, H. Chertkow." The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment," *J. Am. Geriatr. Soc.* Vol. 53, pp. 695–699, 2005.
- [16] V. Hazan, S. Messaoud-Galusi, S. Rosen, S. Nouwens, B. Shakespeare. "Speech perception abilities of adults with dyslexia: is there any evidence for a true deficit?", *Journal of Speech, Hearing and Language Research*, vol. 52, no. 6, pp. 1510–1529, 2009.
- [17] R.S. Bolia, W.T. Nelson, M.A. Ericson and B.D. Simpson. "A speech corpus for multitalker communication research," *Journal of the Acoustical Society of America*, vol. 107, pp. 1065–1066, 2000.
- [18] J. Bench, A. Kowal and J. Bamford. "The BKB (Bamford-Kowal-Bench) Sentence Lists for Partially-Hearing Children," *British Journal of Audiology*, vol. 13, pp. 108–112, 1979.
- [19] D. Wechsler. "WISC-V: Technical and Interpretive Manual," Bloomington, MN: Pearson, 2014.
- [20] D. Wechsler "WAIS-IV administration and scoring manual," San Antonio, TX: Psychological Corporation, 2008.
- [21] Audio3D virtual room software: www.phon.ucl.ac.uk/resource/audio3d/
- [22] Virtual room simulation for Diapix speech recordings: <https://github.com/outepi/Diapix-virtual-room>