STUDY PROTOCOL

The effects of auditory spatial training on informational masking release in elderly listeners: a study protocol for a randomized clinical trial [version 1; peer review: awaiting peer review]

Farnoush Jarollahi¹, Marzieh Amiri ¹, Shohreh Jalaei², Seyyed Jalal Sameni¹

¹Department of Audiology, Faculty of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran
²Department of Physiotherapy, Faculty of Rehabilitation Sciences, Tehran University of Medical Sciences, Tehran, Iran

Abstract

Background: Regarding the strong auditory spatial plasticity capability of the central auditory system and the effect of short-term and long-term rehabilitation programs in elderly people, it seems that an auditory spatial training can help this population in informational masking release and better track speech in noisy environments. The main purposes of this study are developing an informational masking measurement test and an auditory spatial training program.

Protocol: This study will be conducted in two parts. Part 1: develop and determine the validity of an informational masking measurement test by recruiting two groups of young (n=50) and old (n=50) participants with normal hearing who have no difficulty in understanding speech in noisy environments. Part 2 (clinical trial): two groups of 60-75-year-olds with normal hearing, who complain about difficulty in speech perception in noisy environments, will participate as control and intervention groups to examine the effect of auditory spatial training. Intervention: 8 sessions of auditory spatial training. The informational masking measurement test and Speech, Spatial and Qualities of Hearing Scale will be compared before intervention, immediately after intervention, and one month after intervention between the two groups.

Discussion: Since auditory training programs do not deal with informational masking release, an auditory spatial training will be designed, aiming to improve hearing in noisy environments for elderly populations.

Trial registration: Iranian Registry of Clinical Trials (IRCT20190118042404N1) on 25th February 2019.

Keywords
Informational masking, Energetic masking, Elderly, Train, Speech perception in noise
Corresponding author: Marzieh Amiri (amirimarzieh384@gmail.com)

Author roles: Jarollahi F: Conceptualization, Funding Acquisition, Methodology, Supervision, Validation, Writing – Review & Editing; Amiri M: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; Jalaei S: Formal Analysis, Methodology; Sameni SJ: Investigation, Project Administration, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This study was part of a Ph.D. Dissertation approved by Iran University of Medical Sciences (IUMS), Tehran, Iran and is financially supported by IUMS (Contract No: 97-4-6-13657).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2019 Jarollahi F et al. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Jarollahi F, Amiri M, Jalaei S and Sameni SJ. The effects of auditory spatial training on informational masking release in elderly listeners: a study protocol for a randomized clinical trial [version 1; peer review: awaiting peer review]
F1000Research 2019, 8:420 (https://doi.org/10.12688/f1000research.18602.1)

First published: 09 Apr 2019, 8:420 (https://doi.org/10.12688/f1000research.18602.1)
Introduction
Understanding speech in noisy environments is a major challenge of the auditory system, which occurs mostly due to aging. It is clear that poorer speech recognition in the elderly population can occur due to various factors, including peripheral hearing impairment or decline in cognitive capabilities and processing defects at supra threshold levels. It is difficult to determine the exact role of each of these factors in developing speech problems in the elderly. Most elderly individuals complain about the difficulty in understanding speech in noisy environments, despite having normal hearing thresholds. In addition, the type of background noise heavily influences the extent of the damage imposed to speech intelligibility. Generally, hearing problems worsen in noisy environments when the target speech is covered by a competing signal. In this situation, in addition to energetic masking, there is another type of masking known as informational masking.

The spectro-temporal overlap between target and competing speech, which leads to poor target identification, is called energetic masking. Energetic masking is caused by physical interactions between target and competing speech at the low level of the peripheral auditory system. Recent research has suggested that when competing signals occur randomly or when there is a high similarity between target and competing signals (for example, when both signals are speech), another type of masking occurs both. This type of masking, which occurs in response to uncertainty of the competing signal or similarity between target and competing signals, is called non-energetic or informational masking. It leads to failure in selecting auditory objects and therefore impairs auditory scene analysis. Generally, in contrast with the energetic masking, which occurs due to the limitations caused by frequency selectivity at peripheral levels, informational masking reflects the processing capacity limitations at central auditory levels.

Various studies have indicated that with an increase in age, side-effects of competing noises will increase. Different studies have revealed that elderly populations, who do not have peripheral auditory impairment, suffer from diminished ability of using acoustic and phonetic signs to separate speech from background noise, compared to young people; therefore, more informational masking occurs in this population. Since problems with understanding speech will reduce social interactions of the elderly population, it is very important to develop effective auditory rehabilitation programs to prevent their isolation and to improve their quality of life.

Auditory spatial processing plays an important role in speech recognition in complex noisy environments, since it enables the listener to differentiate the target signal from competing signals via auditory scene analysis and forming auditory streams. Based on the results of different studies, it has become clear that the most important sign of informational masking release is spatial separation of target and competing signals. In addition, it has been shown that auditory spatial processing ability is lower in the elderly with normal hearing than in young people. The reduction of localization accuracy and taking advantage of auditory spatial processing, consequently decreasing binaural processing, are not totally related to impaired hearing thresholds. Hence, it leads to poorer speech recognition in elderly people with normal hearing in noisy environments. On the other hand, it has been shown that the elderly population need a higher signal to noise ratio for speech recognition in the presence of noise, compared to young people. These changes are possibly due to the reduction in the ability of using acoustic and phonetic signs to separate target signals from background noise. Therefore, in elderly people, without considering the hearing impairment, the ability to use spatial and non-spatial signs for informational masking release diminishes due to the reduction of cognitive processing abilities, temporal processing defects, defects in the connection between hemispheres, and diminished ability to separate simultaneous sounds.

Current neuroscientific studies have suggested that the central auditory system has a strong neuroplasticity capability for auditory spatial processing and since the effect of short-term and long-term auditory rehabilitation programs has been demonstrated in elderly people, it seems that by providing auditory spatial training, we can aid the elderly population to perform informational masking release, preventing them from missing conversations in noisy environments.

The present study had two parts. The first part of this research was developing a test for measuring and evaluating the informational masking. The second part of this research was a clinical trial of an auditory spatial training program in elderly people with normal hearing, which could diminish informational masking. The main hypothesis for the second part of the study is that presenting an auditory spatial training for elderly people would be effective in the improvement of speech recognition in noisy environments by stimulating the centers related to binaural processing.

Protocol
This is version 1 of this protocol.

Study design, setting and participants
This research will be conducted in the Audiology Clinic of Rehabilitation Faculty of Iran University of Medical Sciences.

This study consists of two main parts. Part 1: develop and determine an informational masking measurement test and explain its validity characteristics in a test development study, conducted cross-sectionally. The study population will be a group of elderly (60 to 75 years old) and a group of young (20 to 40 years old) people. The young people will be recruited from rehabilitation students of Iran University of Medical Sciences, while elderly people are those referred to the audiology clinics of Iran University of Medical Sciences.

Part 2 (simple randomized clinical trial): the effect of training on informational masking release. This part of study is a simple randomized clinical trial design and patients will be randomly assigned into two groups of control (not receiving auditory spatial training) and intervention (receiving auditory spatial
training). The random allocation will be performed based on balanced randomization [1:1] where the allocation will be applied by random number table (those assigned an odd number, control group; those assigned an even number, intervention group). This allocation sequence will be generated by one of the audiology clinic staff of the IUMS who will not have any role in the study. An elderly population, 60–75-years-old, who are referred to the audiology clinics of Iran University of Medical Sciences will be selected. The two groups will be matched for age and gender. Those in the control group will not receive any rehabilitation programs during the study.

**Inclusion criteria (for all participants in the study):** auditory thresholds ≥25dB within the 250–2000Hz frequency range and <40dB at 4000Hz frequency, ensuring lack of salient cognitive problems using Mini Mental State Examination (MMSE)\(^3\); having diploma or higher degree; right-handedness (using Edinburgh handedness inventory); speaking Farsi and being monolingual; complaint about speech in noise perception difficulties (just for those in part 2 of the study); and normal condition of middle ear function.

**Exclusion criteria (for all participants in the study):** unwillingness for participation in each step of research and not meeting inclusion criteria.

**Study procedures**

**Part 1: Developing an informational masking measurement test and determining its validity.** When studying the informational masking, use of the coordinate response measure (CRM) has been frequently been introduced as one of the most popular speech materials for evaluating informational masking\(^2\). In these sentences, the same rigid structure with “Ready [call sign] go to [color] [number] now” format is used. In these sentences, eight call signs, four colors, and eight numbers from 1 to 9 can be used. These sentences will be expressed by speakers of different genders\(^2,4,9\). In the present study, 256 sentences will be created for each speaker (8*4*8). Sentences will be expressed by four speakers (two women and 2 men), providing a total of 1024 sentences. Although CRM stimuli have been initially designed to measure the speech perception in the presence of competing signals, these speech materials provide no contextual information; i.e. predicting the given color or number in the phrases is not possible. This is an important factor in measuring informational masking\(^9\).

Since there is no Persian version of these sentences, this research will prepare the sentences and determine their content and face validity and reliability. Then, after selecting the nouns, colors, and numbers used in the sentence, conforming to the main English version, the prepared sentences will be given to experts in this field (audiologists, speech therapists, and linguistics) to determine the content validity. These experts are the academic members of rehabilitation faculties of IUMS, Tehran University of Medical Sciences (TUMS) and Shahid Beheshti University of Medical Sciences (SBUMS). These experts will be mailed a questionnaire to score the validity items (see Table S1, Extended data).

After selecting the best pattern matching Persian language, based on the model presented for recording the sentences, all sentences will be recorded in a studio with the four speakers. In order to record the sentences, all criteria of the English version including the sampling rate of 44.1 kHz and giving 3s to speakers to produce each sentence will be followed. Then all the sentences will be scaled and all the words in CRM will be set such that they occur simultaneously, called coordinate sentences\(^2\) where each of the sentences will be filtered using a band pass filter of 80 to 8000 Hz filter. Again, in order to determine the face validity, the recorded sentences will be given to the experts mentioned above to determine their suitability. They must fill the questionnaire, which will be emailed to them.

To determine the reliability of CRM speech materials, the mean scores of CRM recognition in the silent will be evaluated in a group of young and old people with normal hearing who do not have speech perception difficulties in noisy environments. There will be one preliminary test and then a re-test. This evaluation will be implemented at the comfort hearing level of the participant. The score will be calculated based on the correct recognition percentage of the sentences. A sentence will receive a correct score when the color-number combination recognition is recognized correctly\(^2,24\). In this study, the mean correct score for color, number, and noun will be studied separately in order to calculate the error percentage for each of them\(^2\). By preparing these sentences, they can be used in the part 2 of the study.

The best way to measure informational masking value is determining the score for speech recognition in the presence of meaningful and meaningless competing noise. To this end, the recognition score for Persian version of CRM speech corpus will be measured under two conditions:

**A. In the presence of meaningful competing noise:** The competing signal is selected from the Persian version of CRM corpus, where the call sign, color and number used in the competing sentence is different from those of the target sentence and it will be expressed by a different speaker. The individual will be trained to pay attention to certain target call signs and ignore other signals\(^2,24\). As one of the important effective factors for informational masking is the great similarity between target and competing signals (like when both of them are speech)\(^6,9\), using CRM sentences as both the target signal and competing signal, the high semantic and syntactic similarity would develop between target and competing signals\(^2,22\).

**B. In the presence of meaningless competing signal:** The previous signal is manipulated such that its spectrum content remains fixed but meaningless - indeed, energetic masking remains but informational masking is reduced. “Time-reversed speech” will be used for this purpose. This is one of the most effective methods in behavioral and neurophysiological research performed for the effects of speech signals on each other. In this method, by fixing the long-term acoustic spectra of two signals and manipulating one of them such that it divides into non-overlapping time segments, and with reverse time
presentation of each segment connecting them to each other, we will have a signal which is equal with the first signal in terms of the spectrum but is not understandable\textsuperscript{30}. In the case of using 20–40 millisecond time windows, this method does not have significant effectiveness in non-understanding the speech signal; therefore, longer time windows should be used\textsuperscript{30}. MATLAB R2018 software will be used in constructing this signal.

The signal to noise ratio in sentence recognition test in steps A and B was ±10, ±5, and 0. The target signal is always presented from a loudspeaker in the 0-azimuth degree and two competing signals from the loudspeakers, which are at ±90 degree and 0 azimuth degree (once with spatial separation and once in the direction of target signal), where once the competing signal has the same gender as the target signal and another time has a different gender. As a result, 20 conditions will be evaluated at each step (5 signal to noise ratios and two spatial angles with two different genders).

Finally, the informational masking score in all 20 conditions will be calculated as follows:

Speech recognition score in the meaningful competing noise condition-speech recognition score at non-understandable noise condition=informational masking score (percentage).

In this step of the research, construct validity will be used to determine the validity of the test. For this purpose, Speech, Spatial, and Qualities of Hearing Scale (SSQ) questionnaire score of each individual will be compared against the informational masking score\textsuperscript{30}. Figure S1 (Extended data) represents the participant’s timeline of the first part of the study.

**Part 2: the effect of auditory spatial training on the informational masking release.** This part of research will be conducted in three steps: before auditory spatial training, during training, and after training.

1. **Assessments prior to auditory spatial training (preliminary interview)**
   - Obtain patient history to confirm the inclusion and exclusion criteria of the participants
   - Initial clinical examination, including otoscopy and tympanometry
   - Perform pure tone audiometry test
   - Perform MMSE questionnaire to ensure lack of salient cognition problem in the participants\textsuperscript{31}
   - Determine speech perception difficulties in the presence of noise: this was evaluated with a question: Do you have difficulty in understanding speech in noisy situations? There were three response options: yes, no or sometimes. Those who responded yes were entered into the study.

   - Measure informational masking score using the test constructed in Part 1 (primary outcome)
   - Determine the SSQ score\textsuperscript{30}. The SSQ self-assessment questionnaire will be filled out by the researcher during the preliminary interview. As improving informational masking can improve the quality of life of people, this questionnaire will be used to measure the quality of life of the participants quantitatively (secondary outcome).

2. **Providing auditory spatial training (intervention group only)**

   Auditory spatial training is designed based on five signs that are important in informational masking release: angular differentiation between target and competing signals\textsuperscript{48,27,20}; signal to noise ratio\textsuperscript{52}; similarity and difference between the target and competing signals\textsuperscript{12,30}; similar or different gender for target and competing signals\textsuperscript{12,30}; and meaningfulness of the competing signal\textsuperscript{52}. Training sessions will be divided into three general steps by considering the competing signals. In the first step, meaningless competing signals will be used like white noise. In the second step, in order to make the training process somewhat difficult, meaning-carrying signals like speech babble consisting of four speakers will be used. Finally, sentence materials with male and female genders will be used. The reason for using the gender factor is that consideration of gender similarity or difference between target and competing signals is one of the signs that adults use for informational masking release\textsuperscript{12,20}.

   In all steps, the target signal will be presented from the loudspeaker at 0-azimuth degree and competing signals from different azimuth angles such as ±90 and 0\textdegree\textsuperscript{11,30}. Therefore, the difficulty of training will grow in each step by reducing the azimuth angle of the competing signal\textsuperscript{30}.

   Sentence signals are selected from Persian version of QuickSIN\textsuperscript{34}. Every step of training is implemented as follows:

   The intensity of the competing signal is fixed at 60 dBsPL, and at the beginning the intensity of target signals is 70 dBsPL. Three first sentences are used for familiarization. If an individual needs more practice, more familiarization sentences are provided.

   An individual is requested to identify the keywords heard in the target sentences. In the case of true and false identification, required feedback is provided. If the individual identifies more than 50% of the keywords, the sentence is considered true. In this signal to noise ratio, 5 sentences are provided where if the individual identifies more than 50% of the presented sentences, the signal to noise ratio decreases in 5dB steps, after which 5 sentences will be provided again for the individual. If the individual does not have the capability to correctly identify more than 50% of the presented sentences in each signal to noise ratio, the training begins where this signal to noise ratio will be considered as the initial level. The training will continue for 20 minutes and the intensity will change in an adaptive procedure, such that in the case of correct identification of a sentence, the intensity will be decreased by 1.5dB while it...
will be increased by 2.5 dB if the individual scored less than 50% of words correctly. At each intensity where an individual can correctly identify the sentences, the next sentence will be presented and the above process will repeat.

The optimal condition for perceptual auditory learning includes active listening to high repetition of signals during the consecutive educational sessions, which is conducted within a short time interval. Since long-term training is not a very suitable option in the clinic, trainings are repeated twice a week over 8 sessions.

**3. Assessments after auditory spatial training (interview immediately after and one month after training)**

The informational masking test (as per Part 1) will be done immediately after training and one month after using the Persian list of the coordinate response measure (CRM) corpus, which will be compared with the pre-training results (preliminary interview). This score will be the primary outcome. The reason for repeating experiments one month after the intervention is determining the reliability of the results obtained by intervention for informational masking release.

The informational masking release value will be calculated based on the difference between sentence recognition score (in all 20 conditions of signal to noise, different spatial angles, and two genders) in both noise situations (meaningful and non-understandable). The changes in informational masking in the assessments will be calculated before and after the intervention across all 20 conditions (see Table S2, Extended data).

As the ultimate purpose of this research is improving the quality of life of elderly people, the score of SSQ immediately and one month after intervention will be obtained and the results of both intervention and control groups will be compared separately. This score will be the secondary outcome of the intervention. Figure S2 (Extended data) represents participants timeline of the second part of the study.

**Sample size**

**Part 1.** The study of Terwee et al. was considered as the basic study to determine the sample size of the first part of our study. They suggested that at least 50 patients in each group must be included to evaluate the construct validity\(^9\). In total, 50 young people aged between 20 and 40 years and 50 elderly people aged between 60 and 75 years, with normal hearing who do not suffer from speech understanding in noisy environments, will be recruited.

**Part 2.** The following formula is used to determine the sample size:

\[
n = \frac{\left( Z_{1-\alpha/2} + Z_{1-\beta} \right)^2 \left( S_1^2 + S_2^2 \right)}{\left( \mu_1 - \mu_2 \right)^2}
\]

- \( S_1 \): standard deviation of the studied variable in the first group (case, exposed, or intervened)
- \( S_2 \): standard deviation of the studied variable in the second group (control, unexposed, or compared)
- \( \mu_1 \): mean of the studied variable in the first group
- \( \mu_2 \): mean of the studied variable in the second group
- \( \alpha = 0.05 \)
- \( \beta = 80\% \)

In this formula, the studied variable is the extent of informational masking changes before and after the intervention. There is no previous study, which was used the same training as this study proposes; therefore, we considered the study of Delphi et al. which was on a group of elderly individuals\(^8\). The sample size of this study was 16 patients in each group. We will use the same sample size.

**Data analysis**

Central tendency and dispersion indices (mean and standard deviation) will be used in descriptive analysis of data. In data analysis and for determining the reliability of the Persian version of the coordinate response measure (CRM) corpus, paired t-test and Pearson correlation will be used in the case of normality of data; otherwise, Spearman test will be employed, and one-way ANOVA will be utilized for determining intra-class correlation coefficient.

In part 2 of the research, repeated measures ANOVA values will be used for inter-group comparisons and two-way ANOVA will be employed for comparison among groups. SPSS software (V20.0, IBM Corporation, New York, USA) will be used for statistical data analysis and the significance level for all tests will be 0.05.

**Ethical statement and consent to participate**

The Medical Ethics Committee of Iran University of medical sciences approved the study protocol (IR.IUMS.REC.1397.303) and the ethical principles of the ethics committee will be observed in this research. Researchers will send any amendments to the protocol in the future to the ethics committee.

One of the researchers of this study will obtain written informed consent from patients willing to participate in the trial (see Extended data). The purpose of the research and its steps will be explained for all participants before the study start. Confidentiality of data and results of tests will be ensured to participants. Participants will be made aware that they can refrain from cooperation in the study when they want. Conducting tests has no side-effects for the studied individuals and all tests and training sessions are without cost to the participants.

To promote participant retention and complete follow-up, in every training session the examiner will provide feedback to all participants and will inform them about the training progress. The researcher will ask them about the impact of training on the participant’s daily communication conversations.
All data will be entered into forms which are prepared for data collection (see Table S2; Extended data) and the participant files will be stored at study site and will be maintained in a secure place and manner. Participant files will be maintained in storage for a period of 2 years after completion of the study. Only Principal Investigators will be given access to the study data.

Dissemination
The study outcomes will be published through peer-reviewed journals. The data resulting from this study will be released to the audiologists and participants and the general medical community. The results of this trial will be communicated to the external funding body through a formal report. There is no limit in the publication of the trial results.

Study status
The study started in December 2018 and will continue until December 2019. To date, the enrolment of the patients has been performed and the allocation will be performed in the near future.

Discussion
Since there is a progressive increase in elderly populations around the world, the independence of this age group has gained much attention, and Iran is no exception. One of the most important points in independent life during aging is the capability for effective verbal communication. Unfortunately, this capability declines in elderly people, especially tracking speech in environments where several speakers talk with each other. Most elderly individuals complain that despite good hearing, they cannot understand speech in noisy environments. Indeed, elderly people cannot use auditory spatial signs for informational masking release due to the reduction of their auditory processing and cognitive abilities. Since informational masking has an important role in competing signal environments and rehabilitation programs have not considered this an important aspect of masking, designing training that can help elderly people in releasing this masking is novel. Therefore, if the main research hypothesis, i.e. auditory spatial training can improve informational masking release in the elderly people, is confirmed, by providing the therapeutic solution in this age group, a series of auditory spatial trainings based on informational masking release will be provided for audiologists. In addition, the Persian version of the coordinate response measure (CRM) corpus and its reliability and validity will be calculated to be used in research on speech recognition in noisy environments.

Data availability
Underlying data
No data is associated with this article.

Extended data

This project contains the following extended data:
- Table S1: The questionnaire of content and face validity of Persian version of coordinate response measure (CRM) corpus
- Table S2: The data collection sheet for the second part of the study
- Figure S1: Participant timeline for part 1 of the study (Developing an informational masking measurement test and determining its validity)
- Figure S2: Participant timeline for part 2 of the study (the effect of auditory spatial training on the informational masking release)
- Informed consent form for the participants of the first part of the study
- Informed consent form for the participants of the second part of the study

Data are available under the terms of the Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).

Reporting guidelines

Grant information
This study was part of a Ph.D. Dissertation approved by Iran University of Medical Sciences (IUMS), Tehran, Iran and is financially supported by IUMS (Contract No: 97-4-6-13657).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements
The authors acknowledge Dr Mahnaz Ahmadi for her good guidance.

References
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com