Investigating the Role of Working Memory in Speech Perception in Noise

Morteza Hamidi Nahrani1* and Abdollah Moossavi2
1Department of Audiology, Faculty of Rehabilitation, Iran
2Department of Otolaryngology, School of Medicine, Iran

Abstract

Background and aim: Optimal speech perception in a noisy environment require sensory and cognitive skills. In the cognitive spectrum, working memory (WM) plays the most effective role in speech perception in noise (SPIN), to which different factors contribute such as age, hearing loss and hearing aids technologies. Recent Finding: in young people with normal hearing, no significant correlation exists between working memory and speech perception in noise, while this correlation increases with age. In people with hearing impairment, the role of working memory in speech in noise is more prominent. Hearing aids technology also affects the relationship between working memory and speech perception due to creating distortion.

Conclusion: Working memory plays an important role in speech recognition. This relation is especially strong in the elderly and hearing-impaired people. Therefore, working memory training can improve speech perception in noise in these individuals. The use of hearing aids in hearing impaired people also depends on the working memory capacity. The greater the capacity of working memory, the hearing aids technology will improve more the speech perception in noise. Finally, working memory training may be used to improve listeners’ speech perception in noise.

Keywords: Working memory; Hearing loss; Speech in noise; Speech perception

Abbreviations: PL: Phonological Loop; VSSP: Visuospatial Sketchpad; EB: Episodic Buffer; PTA: Pure Tone Audiometry; SPIN: Speech Perception in Noise; ELU: Ease of Language Understanding

Introduction

Peripheral hearing is usually measured by pure tone audiometry (PTA). Audiologists often observe patients with similar audiogram function differently in terms of speech perception in adverse hearing conditions, because in addition to hearing ability, the supra-threshold differentiation and cognitive abilities play a significant role in the SPIN [1]. In recent years, there has been a growing interest in the role of individual cognitive performance on language and speech processing. In the cognitive performance spectrum, attention skills and WM play a significant role in following SPIN [2]. Studies show that musicians have a higher capacity of auditory WM and perform better in speech tests in noise [3,4] therefore, the relationship between WM and SPIN seems logical and the higher the WM capacity, the better the SPIN [5,6].

Working memory

The term WM was introduced in 1968 by Atkinson and Shiffrin to describe a short-term buffer for storing and processing lingual-verbal-audio information. Sometimes the term “WM” and “short-term memory” are used interchangeably. Short-term memory is in fact part of the WM used to store the data without manipulation, while WM includes storing and manipulating information for mental activities [6]. For a large part of human characteristics, WM gives meaning to verbal and written communication, storing information in mind and matching them over time, decision making and reasoning [7]. Planning for temporary storage and real-time processing of information is the task of WM [8]. Several models and hypotheses have been proposed for WM. Meanwhile, the multi-component model of Baddeley and Hitch has attracted more attention. This model includes the central executive and its three slave systems: the phonological loop (PL), visuospatial sketchpad (VSSP), and the episodic buffer (EB).
PL is for preserving, rehearsing and manipulating phonological information. The VSSP has two subsets: visual storage that stores and manipulates the physical properties of objects such as colour and shape; the space mechanism that helps with motor movements such as gesture and balance movements. EB is a system with limited storage capacity and is able to hold coherent chunks of information in a multi-dimensional code. EB integrates the information of both PL and VSSP slave systems [9]. Central executive is the most important part of WM and it is responsible for controlling attention and solving problems. This part coordinates memory functions and is responsible for retrieving information from the long-term memory storage; it is also responsible for the initial input of information into the processing system. By blocking attention to target processing stimulus, CE immediately suppresses irrelevant information to the processing target [7,9]. Cognitive control, which is a sort of inhibitory function, enables the individual to pay selective attention to the target stimulus and inhibit inappropriate and irrelevant information. It is generally known as executive performance and is important in empowering individuals to control the attention, thoughts, and various activities especially when it is impossible to rely on bottom-up information. This inhibitory function can facilitate the function of WM and increase focus on target stimuli [10]. Eventually, PL and VSSP finalize the information and do not allow this inappropriate information to be coded, maintained and rehearsed [6-9,11].

In short, the four components of WM are likely to act as a dynamic mesh. The mesh allows the entry of desirable information to the cognitive processing system and removes the information undesirable for speech performance from processing. The mesh develops over time in children so that the higher the refining power of the system, the mesh holes become smaller; while in older people, these holes become wider and allow the entry of undesirable information for processing [8].

**Basics of speech perception**

Hearing and perceiving the speech of a compound task involves a wide range of sensory and cognitive processing. Due to the acoustic complexity, speech is susceptible to be covered by other environmental sounds. Therefore, people especially the children and the elderly, are often faced with situations that challenge speech perception in the presence of background noise [12,13]. Effective factors on SPIN are still not well determined. Studies have shown for speech perception in the presence of background noise, both bottom-up audio processing of speech signal and the top-down cognitive and linguistic processing are involved. Lingual and cognitive processing involves long-term linguistic knowledge such as vocabulary and grammar, and WM mechanisms such as short-term phonological memory [14]. From a cognitive point of view, during speech perception listeners should extract and store meaning from acoustic patterns to be used for summarizing with audio stream. When the acoustic patterns are degraded or altered contrary to the expected form, the adaptation of these acoustic patterns with the stored vocabulary information becomes difficult and more WM is used with more capacity [15].

Therefore, the ability to hear a successful speech in the presence of environmental noise requires sensory and cognitive skills. From a sensory point of view, the hearing apparatus should be locked with a target speech signal and eliminate the background noise and competitive sounds. The relative stability of voice pitch or base frequency during a span of time creates an identity in the speech stream and helps separate grouping of speech and noise. Of course, other signs such as time, harmonics, and location contribute to the grouping of speech information that can be reduced in hearing impaired people [16]. Children show more vulnerability to noise for speech perception. Studies show that with the increase of noise level in educational settings, the academic performance significantly decreases. Some children seem to have more difficulty in SPIN than others. In the absence of peripheral hearing loss, these speech perception problems are caused by disorders in the central levels [17]. During the recognition of words of sentences in the presence of noise, children use the semantics of sentences less than adults. Also, the child’s vocabulary does not play a significant role in SPIN [14].

People with Sensorineural hearing loss suffer from disorganized auditory input. Distorted hearing signals may not be consistent with long-term memory phonological representations [18]; therefore, hearing loss can lead to mismatch of input signals with long-term memory representation: the distorted input signal caused by cochlear damage and degraded phonological representation in long-term memory. In these cases, it is necessary to use explicit processing in WM. On the other hand, providing speech signal with noise creates the third source of mismatch in people with hearing loss. For this reason, hearing in the background noise is more difficult for hearing impaired people [19].

**The relationship between WM and SPIN**

Regarding the relationship between WM and speech perception, two important hypotheses have been discussed in the literature: The first is the capacity hypothesis introduced by Just 1992. Based on this hypothesis, WM does a limited amount of activity that is distributed among all cognitive tasks. Memory limitation is due to the fact that processing requires a lot of effort or more time for mental action. The more difficult processing is, the slower the system works. When the capacity of WM is maximized, access to other cognitive resources decreases. The functional disorder of the hearing apparatus increases the auditory effort and, as a result, reduces the speed of cognitive processing. In order to assign the retrieval and storage of resources to new tasks, the amount of activity of WM should be reduced [20].

The second is the ease of language understanding (ELU) model. Based on this model, any mismatch between speech input and phonological representation stored in long-term memory disrupts automatic vocabulary retrieval and leads to the use of busy processing mechanisms such as WM. Internal distortions such as problems of hearing apparatus and cognitive function as well as external distortions such as background noise can cause this mismatch.
The second is the ease of language understanding (ELU) model. Based on this model, any mismatch between speech input and phonological representation stored in long-term memory disrupts automatic vocabulary retrieval and leads to the use of busy processing mechanisms such as WM. Internal distortions such as problems of hearing apparatus and cognitive function as well as external distortions such as background noise can cause this mismatch [18]. According to ELU model, the level of cognitive function interference in speech recognition depends on hearing conditions, and the individual cognitive ability can predict the degree of speech recognition in adverse hearing conditions [21]. When hearing conditions are favorable, speech inputs are consistent with phonological representation in long-term memory. This type of processing is automatic and implicit. When hearing is challenging or there is peripheral hearing loss, a mismatch between speech inputs and long-term memory information emerges. In such conditions, explicit processing is required to match the degraded input and representation at the long-term memory storage. For explicit processing, high capacity of WM is required. If the listener has a low capacity in WM, the processing speed of the system is greatly reduced [18,22].

The relationship between WM and SPIN in subjects with normal hearing

Generally, there is no correlation between WM and SPIN [14,23-25]. In noisy environments, people with normal hearing rely mostly on their auditory abilities rather than cognitive abilities [23] and WM capacity cannot be a strong predictor of SPIN [15,24,25]. Unlike the overall prediction of ELU model, no significant relationship was found between WM capacity and SPIN in the children [14] and the elderly [26] with normal hearing. The Millman study, however, contradicted the above results. According to this study, as the scores of WM gets higher, the score of SPIN increases in old people with normal hearing [27]. This inconsistency among the studies seems to be due to the type of noise used in SPIN tests and the type of tests used to check the WM. In Millman study, modulated noise and non-word repetition (NWR) test have been used. Moreover, Mc Creery et al. [27] showed a significant correlation between WM and sentence perception in noise, while there was no correlation for perception of one-syllable words. In the study of Jessica, children with normal hearing experienced a decrease in auditory WM at a signal-to-noise ratio of 0 and -5 DB. [28,29]. Finally, there was no consistency between the results of various studies on the relationship between WM and SPIN in normal hearing subjects; this inconsistency could be also due to using different tests for evaluating WM and the difference in the type of target stimulus and competitive signal used in SPIN test.

The role of age on the WM capacity

The study by Fullgrabe and Rosen et al. [24] suggests that with increasing age, the role of WM increases in speech perception in adverse listening conditions. The probable cause of this is the accumulation of age-related disorders and auditory processing disorders in the temporal-fine-structure and temporal envelope cues that result in insufficient internal representation of speech, and therefore calls for compensatory mechanisms such as WM to help speech perception [15]. The results of studies indicate that the role of WM capacity on SPIN cannot be evaluated independently of the age and health of the hearing apparatus [15,30].

Gordon Salant et al. [22] showed that elderly adults with high WM capacity act similarly as young people in recognizing sentence in noise. Also, speech recognition in noise is difficult for the people with low-capacity WM and normal hearing whether young or old. Generally, the WM capacity is reduced with age, but there is a high degree of variation in all age groups.

The relationship between WM and SPIN in hearing impaired people

Hearing loss has a negative effect on verbal communication, especially in noisy environments. Hearing impaired children have weak hearing skills, but their visual WM has a good performance. Therefore, this weakness in verbal auditory memory can reduce SPIN [31]. The results show when hearing loss develops, non-verbal cognitive ability decreases [32]; while in the study of Olson and Campbell, fewer errors occurred in the reading span test by the increase of hearing loss. According to this study, the elderly with moderate-severe hearing loss show better performance in WM tests than those with mild hearing loss. The duration of hearing loss in the moderate-severe hearing loss group was twice as the duration of mild hearing loss; people with a longer duration of hearing loss adjusted with hearing impairment. Besides, in severe hearing loss, auditory cortex areas respond to visual stimulus, which can lead to improvement of R-Span test [33]. According to Nagaraj, there is no correlation between the amount of hearing loss and WM; therefore, hearing impairment is not an interfering factor in assessing the amount of WM, and auditory WM tests can be used to assess hearing impaired people [34]. Finally, the role of hearing loss in WM can be justified by the ELU model. In the presence of hearing loss, speech input to brain is degraded and creates mismatch between speech inputs and information stored in memory. A listener with high-capacity WM can compensate for this mismatch and maintain SPIN.

The role of hearing aids on the relationship between WM and SPIN

The ability to communicate in adverse listening conditions is reduced in people with hearing impairment. Hearing aid is recommended for improving communication in hearing impaired patients [35]. A greater WM capacity leads to better speech recognition in noise with and without the hearing aids [36]. In addition, hearing aid users with high WM capacity are less affected by signal distortion due to signal compression [19,37]. The use of hearing aids with adequate amplification and proper signal processing can lead to more speech clarity in people with hearing loss. According to ELU model, this improvement in speech clarity is expressed as a reduction in mismatch and a consistent reduction in explicit processing interference, therefore more resources will be available for cognitive processing. Of course, the processing of a hearing aid signal may have adverse effects on speech input and lead to mismatch. The longer the hearing aid is used, this mismatch
is reduced more since a new phonological representation is formed in long-term memory [38,39]. To reduce mismatch and create a new representation, use of hearing aid is required for 4-9 weeks of habituation [40]. Rudner evaluated speech recognition function in two types of settings. When a person with hearing impairment was not familiar with hearing aids, the WM function played a more important role in speech recognition, and the effect was reduced in a familiar setting to the patient [41]; therefore, with greater habituation of the patient more cognitive skills become available to improve SPIN.

NG’s study on hearing aid users showed that the relationship between WM and the recognition threshold of speech at the onset of using hearing aids was the strongest and, over time, the relationship reduced. Immediately after wearing a hearing aid, the reading span test was the main predictor of speech recognition function, while six months after receiving a hearing aid, the mean hearing threshold was a stronger predictor of speech perception. This means that experienced users of hearing aids are less reliant on WM to recognize speech, so explicit processing is reduced. Assuming that cognitive module is fixed, other cognitive resources can be easily used to understand speech. Therefore, it is completely logical to suggest that speech perception in users of hearing aids is improved over time [38]. From the theoretical point of view, fast compression in WDR circuit, which is more beneficial for brief speech segments, provides more capability for consonant intelligibility. Of course, the altered envelope caused by fast compression can alter the speech signal and, ultimately, leads to adverse listening conditions for listeners who rely on envelope cues for hearing [42].

Studies show in hearing aid users with high-capacity WM who have a better hearing ability at inter-noise intervals, compression with short release time leads to a higher speech perception, while in people with lower cognitive ability compression with long release time is suggested [40,43-46]. Another signal processing that leads to signal changes is frequency compression. In frequency compression, the energy of the signal at high frequencies is digitally transmitted to the adjacent lower frequency region. The aim of frequency compression is to improve hearing, but at the same time, by shifting the harmonics and changing the levels of peaks of the spectrum, it results in change of sound quality, making speech perception difficult for people with low WM capacity. The relationship between WM and frequency compression processing is complex and may be influenced by various factors [47]. Therefore, the role of frequency compression in auditory improvement depends on the WM capacity, type of processing frequency compression and the amount of hearing loss [19,48].

The main purpose of the noise reduction circuit is to reduce the adverse effects of background noise on the speech. In this regard, 8 studies have been carried out, in four of which the relationship between WM and noise reduction processing is reported. As the WM capacity is higher, noise reduction circuit is more influential on SPIN [49-52]. There was no such relationship in the other four studies [53-56]. Due to these divergent findings, these findings could not be adjusted with ELU model; the reason may be that noise reduction processing has two simultaneous effects: improved hearing of the signal in noise from one side and creating a larger distortion due to the performance of this processing; the combined effects of these two factors are likely to be the cause of the weak relation between WM and the processing of reduced noise.

According to Rudner, hearing aid users with a good WM capacity benefit more from signal processing because they have access to more cognitive resources to overcome the distortions caused by signal processing [40]. As a result, any change in the output of the speech signal by the hearing aid can lead to speech perception vulnerability in hearing impaired people with low WM capacity; therefore, the role of hearing aids technology on improving SPIN is bilateral. These technologies, on the one hand, provide the ability to hear and, on the other hand, cause distortion in speech inputs, which are reduced over time [42].

The role of training and rehabilitation of WM on SPIN

As mentioned above, in some groups there is a correlation between WM capacity and SPIN. Therefore, if WM exercises increase the WM capacity, it is likely to improve SPIN. Various studies have been done on WM training; however, the effectiveness of WM training and general cognitive education has been doubted up to day [57,58]. In Wayne RV et al. [57] WM training was studied in old people with normal hearing using Cogmed software. In this exercise, which was conducted as 25 daily sessions at home without supervision, a slight improvement in verbal WM was observed, but the performance of SPIN did not show an improvement [59]. Similar results were observed in the study of Rudner M et al. [58]. These results contradict the studies by Ingvalson et al. [59] on young people with normal hearing and Ashwini on older people with normal hearing, which reported a significant improvement in the performance of WM and SPIN after receiving WM training [60].

Several studies have also been performed on cochlear implant users Kronenberger et al. [61] studied young children with cochlear implantation and reported short-term improvement in verbal and nonverbal WM capacity and sentence perception after 5 weeks of training through Cogmed software; but during 6-months follow-up, this improvement in WM and sentence perception was reduced, in a way that did not significantly differ with pre-training assessments. In the study of Doosti et al. [62] significant improvements were observed in WM capacity of 25 Iranian children with cochlear implant using Cogmed software. In this study, the speech perception was not studied. It seems that differences in the type of training, the number of training sessions, the evaluated groups and absence of a control group in some studies have contributed to this variation in the effectiveness of WM training.

Conclusion and Clinical Application

The results on the correlation between WM and SPIN show varying results in different subjects. According to ELU theory, the greater the auditory inputs are degraded and distorted, the role of WM in speech perception becomes more important; therefore, in hearing impaired individuals, there is a greater correlation between WM capacity and SPIN, and WM is one of the important
components of speech perception and communication in hearing impaired people. With regard to the above, the evaluation of WM and SPIN in people with hearing impairment seems logical. In hearing aids users who have a low capacity of WM, the use of fast compression can lead to degrading SPIN; while in people with high capacity WM fast compression can facilitate SPIN. Of course, the use of other hearing aids technologies, such as frequency compression and noise reduction, is still under debate. In general, the use of signal processing in hearing aids for users may depend on the cognitive ability of the patient, and cognitive assessments are necessary before wearing the hearing aids.

Persons with a higher capacity of WM show better performance in speech recognition in noise tests. Therefore, if WM training can increase WM capacity, the challenge of SPIN can be reduced, especially in the elderly. Various methods have been proposed for WM training. Amongst all, cognitive software has become more popular. This software covers all parts of the WM and is run as a game. This software is increasingly expanding among professionals as a multi-million-dollar industry. However, conflicting results on the effectiveness of this rehabilitation software on improving WM have been reported. Still, no comprehensive has been reported which is conducted in target groups such as hearing-impaired subjects with a high sample size. It is suggested that to investigate the effectiveness of these software in hearing impaired people with a large sample size and cognitive assessments in future studies. If WM training contributes to SPIN, these trainings should be considered as one of the main rehabilitation programs for the hearing impaired and elderly people with SPIN problems.

References


For possible submissions Click below:

Submit Article