Int. Adv. Otol. 2010; 6:(2) 258-262

REVIEW

Time Course of Auditory Processing, Visual Processing, Language and Speech Processing

Nuray Bayar Muluk, Fulya Yalcinkaya

Kırıkkale University, Faculty of Medicine, ENT Department, Kırıkkale, Turkey, (NM)
Hacettepe University, Faculty of Medicine, ENT Department, Division of Audiology and Speech Pathology, Ankara, Turkey, (FY)

Each stimulus is processed in the brain at a certain speed/time. Hearing, vision and language are included in this process. Such as, the onset of language specific phonetic–phonological analysis has been estimated at 100–200 ms. Listener the smaller the gap that can be detected. Such as, rapidly changing (gap) sounds such as /r/, /l/. There are need both short (20msec for phoneme duration signals) and long (200msec for syllable-duration signals) segments of speech.

In hearing, language and speech processing functions, brain works together with all fields (auditory processing, memory, language and the image and speech recording area, etc.) synchronizely for seconds as the orchestra. If neurons can not participate this processing synchronizely, synchronization is corrupted. Processing time of information and synchronization work should be the basis for hearing, language and speech training. Phonetics in speech come to our ears in a few seconds through sound waves. If these sounds can not received within a few seconds, they get lost. If received, they were processed in the auditory pathway and brain in a few seconds. The purpose of this review is to draw attention that, if the sounds are received and processed within a few seconds, the training method used in speech training model should be intended for sounds' transmission and processing in a few seconds. In addition, all of the functions (auditory, view processing, memory and language) should be included into training by bottom-up approach.

Auditory processing is the ability to listen, comprehend and respond to information that we hear through our auditory channels. It needs decoding of the neural message. Auditory processing involves attention to detection and identification of the signal; and decoding of the neural message. If we don't give full attention to heard things, listening difficulty ocuurs. In poor attention and listening conditions, rapid acoustic changes in speech can not be discriminated.

Submitted : 18 July 2009 *Revised : 16 February 2010* *Accepted : 16 February 2010*

Auditory processing is the ability to listen, comprehend and respond to information that we hear through our auditory channels. It needs decoding of the neural message. Auditory processing involves attention to detection and identification of the signal; and decoding of the neural message. If we don't give full attention to heard things, listening difficulty ocuurs. In poor attention and listening conditions, rapid acoustic changes in speech can not be discriminated.

Auditory Attention

Auditory attention allows us to rapidly and precisely direct our acoustic searchlight toward sounds of interest in our acoustic environment. Speech perception has only recently begun to receive the attention it deserves [1]. In general, auditory attention is thought as only at auditory region. However, in the processing of the stimuli, visual attention also play an important role as auditory attention. Auditory attention as indicated by the visual cue around the fixation cross, and to respond to target duration changes occurring among the attended sounds. To heard sounds clearly, auditory attention is necessary. If full attention can not be given, sounds are heard as muffled. In this case, we mention poor auditory closure and receptive language disorders.

Attention can be top-down (voluntary or task-dependent) or bottom-up (sound-based salience) [1]. For bottom-up attention, auditory input and processing are required. For bottom-up processing, sound conductive mechanism, involving outer and middle ear; and sensorineural mechanism, including inner ear and the auditory nerve should function normally [2,3].

The Central Auditory System

The central auditory system occurs after the ear receives the signal and starts to transfer the stimulus from the inner ear to the cortex via the eighth cranial nevre. Auditory nerve is considered the start of the

Corresponding address:
Nuray Bayar Muluk
Birlik Mahallesi, Zirvekent 2. Etap Sitesi, C-3 blok, No: 62/43 06610 Çankaya / ANKARA-TURKEY
Phone: +90 312 4964073 , +90 532 7182441; Fax: +90 318 2252819; E-mail: nbayarmuluk@yahoo.com • nurayb@hotmail.com

central auditory nervous system. The eighth cranial or auditory nerve runs through Heschl's gyrus in the cortex of the brain. Auditory Nerve Travels from cochlea through internal auditory meatus to skull cavity and brain stem, carry signals from cochlea to primary auditory cortex, with continuous processing along the way[2,3].

The central/auditory processing system includes: 1. ventral cochlear nuclei, 2. dorsal cochlear nuclei, 3. cochlear nuclei, 4. trapezoid body, 5. superior olivary complex, 6.lateral lemniscus, 7.commisure of inferior colluculus, 8. inferior colliculi, 9.medial geniculate body, 10. medial geniculate body, 11. auditory radiation, 12. primary auditory cortex, 13. auditory cortex, 14. corpus callosum. Bottom-up processing is performed at initial and higher levels (Table 1)[4].

There can be a distinction drawn between language processing and central auditory processing. Language processing is mainly speaking of structures in the cortex that are above the brainstem, beginning post Heschl's gyrus. Auditory processing is from the eighth cranial nerve to Heschl's gyrus[4].

"Bottom-up" and "Top-down" Process

Speech comprehension includes both bottom-up and top-down processes[5]. "Bottom-up" is to process the acoustic information from "bottom" to "up". It is used to learn new information. "Top-down" is to give information from "top" to "down". By simple definition, to say the learned information when asked. Addition of new information to learned knowledge; and learning the foreign language are also "bottom-up" processes (Table 1)[4].

Bottom-up process is to take the heard sounds from bottom to up for processing. The process of first the auditory input, such as speech depends of ağabeylity to discriminate changes in the intensity, frequency, and duration of the sounds or-in psychological terms, changes in loudness, pitch and timing[6]. Brain (Auditory Cortex), Wernicke's Area within temporal lobe, save these sounds and its mean; and record this information to language related processing centers.The frontal and temporal areas in speech comprehension in that temporal regions subserve bottom-up processing of speech, whereas frontal areas are more involved in top-down supplementary mechanisms.

Importance and support of visual information

Auditory attention as indicated by the visual cue (beat gestures, head nods, eyebrow movements) around the fixation cross, and to respond to target duration changes occurring among the attended sounds are important. Cortical correlates of seen speech suggest that at the neurological as well as the perceptual level, auditory processing of speech is affected by vision, so that 'auditory speech regions' are activated by seen speech.

Cortical correlates of seen speech suggest that at the neurological as well as the perceptual level, auditory processing of speech is affected by vision, so that 'auditory speech regions' are activated by seen speech. The processing of natural speech, whether it is heard,

Table 1. The Auditory System (Initial and Higher Levels); and Bottom-up or Top-down Processing Schema (4)

			Processing	
Auditory System*	**Anatomic Landmarks***	**Physical World**	**Bottom-Up**	**Top-down**
Initial Auditory System (Peripheral System)	External and middle ear Basilar membrane Inner hair cell Inner ear 1st auditory nerve	Auditory Processing	↓	↑
Central System (Brainstem)	Cochlear nerve and cochlear nucleus Superior olivary complex Lateral lemniscus Inferior colliculi Medial geniculate body		↓	↑
Higher level Auditory System	Brain Cortex (Auditory Cortex, Association Cortex)			
		Cognitive & Mental World		

seen or heard and seen, activates the perisylvian language regions (left>right) [7].The amount and nature of visual information extracted during this period is proposed to initiate the speech-processing system, in which the formation of an abstract representation is updated continuously through visual inputs, up to the point of explicitly registering auditory input.

The effects of visual beats on prosodic prominence: effect on acoustic analyses, auditory perception and visual perception [8]. Corbetta and Shulman [9] suggested that partially segregated areas of the cerebral cortex are involved in “bottom-up” triggered (or stimulus-driven) and top-down controlled shifts of visual attention. Speech comprehension includes both bottom-up and top-down processes [5]. The study, they found distinct roles. The frontal and temporal areas in speech comprehension in that temporal regions subserve bottom-up processing of speech, whereas frontal areas are more involved in top-down supplementary mechanisms.

"What" and "Where" Working Memory in Processing Sounds

In the processing of natural speech, it is important to determine where the sound is coming from. Goal-directed attention to sound identity (what) and sound location (where) has been associated with increased neural activity in ventral and dorsal brain regions, respectively. Relation (Spatiotemporal Analysis of Auditory "What" and "Where" Working Memory study results show that for identical sounds, top-down effects on processing "what" and "where" information is observable at about 200 ms after sound onset and involves a widely distributed neural network [10]

The amount and nature of visual information extracted during this period is proposed to initiate the speech-processing system, in which the formation of an abstract representation is updated continuously through visual inputs, up to the point of explicitly registering auditory input. Bottom-up processing and auditory attention, hearing beats, visual leads make auditory perception easier.

One’s ability to perceive auditory objects in everyday acoustic environments depends on localization and identification of relevant sounds. Presumably, the human auditory cortex “what” and “where” pathways interact closely to facilitate perception of auditory objects. Notably, this differential “what” vs. “where” response adaptation occurred ≈100 ms after sound onset [11]

Generally speaking, speech perception proceeds through a series of stages in which acoustic cues are extracted and stored in sensory memory and then mapped onto linguistic information. We can select the information that is most important to us given moment.

Perceptual acoustic storage and echoic memory

In Bottom-up processing, auditory memory is important for auditory perception. Sound, heard at the instant time, must be received by short term (echoic memory) memory. Individual’s ability to extract featural information increases steadily up to 250 miliseconds after onset [12]. Recognition of an auditory stimulus is not immediate, it must somehow be stored; this is thought to occur in perceptual acoustic storage, often echoic memory. Echoic memory/ Feature extraction [13], audibility of acoustic feature:

a) Temporal resolutions (minimum time interval required to segregate of resolve events. Listener the smaller the gap that can be detected). Such as, rapidly changing sounds such as /k/, /g/.
b) Frequency resolution (to discriminate successive sounds on the basis of frequency).
c) Intensity/phase of signals reaching the ears (interaural aural diffrences)(noisy situations: target signal from the competing background.

Features of sounds such as time and intensity are important binaural cues for localizing their sources. Interaural time differences (ITDs) and interaural level differences are extracted and processed in parallel by separate pathways in the brainstem auditory nuclei [14]. During auditory language comprehension, listeners need to rapidly extract meaning from the continuous speech-stream. Such as, by tracking the neural time course of syllable processing with magnetoencephalography, they show that this continuous construction of meaning-based representations is aided by both top-down (context-based) expectations and bottom-up (acoustic–phonetic) cues in the speech signal. Syllables elicited a sustained response at 200–600 ms

(N400m) which became most similar to that evoked by words when the expectation for meaningful speech was increased by presenting the syllables among words and sentences or using sentence-initial syllables. This word-like cortical processing of meaningless syllables emerged at the build-up of the N400m response, 200–300 ms after speech onset, during the transition from perceptual to lexical–semantic analysis. These findings show that the efficiency of meaning-based analysis of speech is subserved by a cortical system finely tuned to lexically relevant acoustic–phonetic and contextual cues [15].

Krahmer and Swerts [8] found that visual beats have a significant effect on the perceived prominence of the target words Short–Term Memory (STM), conscious, brief retention of information that is currently being processed in a child's mind. STM, is a component of the information-processing system in which new information must remain for a minimum of 20-30 s or the information will be lost. Retention of information that undergoes little processing or interpretation and can be recalled for only a few seconds. Short-term memory can retain about seven items. Determined by adjusting the cue given. Depends on what you are looking at. Brightness affects the decay of iconic memory. Iconic memory, the majority of information decays between 300 and 500 msec. Is all the information lost? Memory traces do last than 1 sec. Only if “recognition " is tested. Information is available only for 1 sec under "full report or partial report" conditions. Iconic memory keeps our consciousness from being overwhelmed.

We can select the information that is most important to us given moment. Iconic memory is short lived. Visual memory can be held longer (recognition). Echoic memory is both similar and different to iconic memory. According to visual memory, within the first few seconds of visual memory a great deal of information is lost. The time, in which the information is held, can be influenced by a number of things. However, enough information can be held to make recognition out to a number of days [16, 17].

Such as, when language is the sound being processed, the formants are mapped onto phonemes, which are the smallest unit of sound in a language. For example, in English the phonemes in the word "glad" are /g/, /l/, /æ/, and /d/. (phone: unit of a single identified speech sound description of the production, transmission and perception of speech sounds). In processing of auditory phonetics features of speech sounds (segmental and suprasegmental), procedure include speech sounds of the word “glad”, meaning, when and where to use, which sounds to use after which; and where to store in the memory. By including all sensories in this process, transfer to memory is done; and when these information was called back, new information may be added onto it.

Recognition of an auditory stimulus

A part of the auditory system automatically detects changes in the acoustic environment. The sensory systems are normally ascribed a role in the continuous, moment-to-moment updating of a veridical mental representation of the stimulus as it unfolds in time [18]. Recognition of an auditory stimulus is not immediate, it must somehow be stored; this is thought to occur in perceptual acoustic storage, often echoic memory. The impotance of the echoic memory stage is that it enables the listener to continue processing a signal no longer in physical existence. Because, an individual’s ability to extract featural information increases steadily up to 250 miliseconds ofter onset. It was found that attentional mechanism do not affect performance in the echoic memory stage. However, attention does greatly influence performance for synthesized auditory memory and subsequent stage of processing. That is, for intervals greater than 250 miliseconds signal offset, attention can significantly enhance the processing of target stimuli [12].

Conclusion

As seen in the above, brain works with all senses synchronizely as orchestra. Therefore, all the senses should be included into hearing, language and speech training; and training methods should be intended for speed and duration. Additionally, training for any processing disorders should not be given only to problematic area. For example, only auditory training for hearing loss; only auditory training for auditory processing disorders; only phonologic and articulation training for speech sound disorders are not right approaches. Therefore, each area of language and speech should be get matured and worked; so that there will not be any delay in participating to synchronized process.

This rewiew may be useful to remind that it is not enough to work towards disorder; because systems are working together and time is very important factor for processing. In central pathways, to process the sounds as bottom-up, auditory attention, visual clue and time are very important.

References

1. Fritz JB, Elhilali M, David SV, Shamma SA. Auditory attention —focusing the searchlight on sound. Current Opinion in Neurobiology 2007; 17: 437–455.

2. Katz J, editor. Handbook of clinical audiology. 5th edition. Baltimore: Lippincott Williams & Wilkins, 2002.

3. Akyıldız N, editor. Ear Diseases and Microsurgery (Kulak Hastalıkları ve Mikrocerrahisi). Volume 1. Ankara: Scientific Medical Bookstore (Bilimsel Tıp Kitabevi), 2002.

4.Richard GJ, editor. The source for processing disorders. East Moline: LinguiSystems, 2001.

5. Zekveld AA, Heslenfeld DJ, Festen JM, Schoonhoven R. Top–down and bottom–up processes in speech comprehension. NeuroImage 2006; 32: 1826-1836.

6. Mauk MD, Buonomano DV. The neural basis of temporal processing. Annual Review of Neuroscience 2004; 27: 307–407.Jääskeläinen IP, Ahveninen J, Bonmassar G, Dale AM, Ilmoniemi RJ, Levänen S, Lin FH, May P, Melcher J, Stufflebeam S, et al. Task-modulated “what” and “where” pathways in human auditory cortex. Proc Natl Acad Sci 2004; 101: 6809–6814.

8.Krahmer E, Swerts M. The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception, Journal of Memory and Language 2007; 57: 396-414.

9. Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain, Nat. Rev. Neurosci. 2002; 3: 201–215.

10. Alain C, McDonald KL, Kovacevic N, McIntosh AR. Spatiotemporal analysis of auditory "what" and "where" working memory. Cereb Cortex. 2009; 19:305-14.

11.Ahveninen J, Jaaskelainen IP, Raij T, Bonmassar G, Devore S, Hamalainen M, Levanen S, Lin FH, Sams N, Shinn-Cunningham BG, Witsel T, Belliveau JW. Task-modulated "what" and "where" pathways in human auditory cortex. Proceedings of the National Academy of Sciences 2006; 103: 1408-1413.

12 Bellis TJ, editor. Assessment and management of central auditory processing disorders in the educational setting from science to practice. Second Edition. Clifton Park, NY: Thomson/Delmar Learning, 2003.

13.Medwetsky L. Memory and attention processing deficits. In M.G. Masters, N. Stecker and J. Katz (Eds.) Central Auditory Processing Disorders: Mostly Management, Boston: Allyn & Bacon, 1998; 63-88.

14. Bonte M, Parviainen T, Hytönen K, Salmelin R. Time course of top-down and bottom-up influences on syllable processing in the auditory cortex. Cereb Cortex 2006; 16: 115-23.

15. Bedny M, Hulbert JC, Thompson-Schill SL. Understanding words in context: the role of Broca's area in word comprehension. Brain Res. 2006; 1146: 101–114.

16. Campbell AJ, Mewhort DJK. On familiarity effects in visual information processing. Canadian Journal of Psychology 1980; 34: 134-154.

17. Chow SL. Iconic store and partial report. Memory & Cognition 1985; 13: 256-264.

18. Phillips DP, Susan E, Hall SE. Psychophysical evidence for adaptation of central auditory processors for interaural differences in time and level. Hearing Research 2005; 202 : 188-199.