VIDEO-METRICS: CONDITIONED HEAD-TURN TECHNIQUES AND BEYOND

Barry A Hoy, Saint Leo University, Eleanor L. Hoy, Norfolk State University

Abstract

The use of the Head-Turn Technique (HTT) as a tool to measure human response to audio cues has been part of psychological research for some twenty years. The tool measures the participant's response to subtle changes in phonetics he or she is hearing by sensing the shift of the participant's attention focus toward the source of the phonetic stimulus. The existing tool has been largely unchanged since its inception. Previous proposals by this researcher presented enhancements to the test procedure; however, the methodology made use of rather primitive and now outdated technology. The tool, as it is, relies heavily upon human interaction, which may be contributing to inaccuracy of measurement and limitations in the types and richness of data that are captured.

In the existing process, a test administrator manually initiates the event prompting the change in the focus of the subject's attention. The occurrence or non-occurrence of a response is then judged by the test administrator. Computer control is limited to the generation of the phonetic stimulus. The proposed revision makes use of Video Image Capture (VIC) and video-metrics to provide more positive and more precise measurements. The enhancement will add the ability not only to detect the occurrence of the head turn event but also to time various aspects of the event. Computer software which is part of the project would track changes in gaze focus with millisecond accuracy. It could also measure the divergence between the orientation of the head and the focus of attention.

Discussion of the Head-Turn Technique in its Present Form

The present iteration of the HTT test process was developed for use in measuring infant and toddler responses to subtle changes in the phonetic composition of sounds to which the child is exposed [1], [2]. Figure 1 presents a plan view of the existing test layout showing one of the team members and the test participant. Figure 1 does not show the second test team member since that member is not present at the test location but administers the test from a remote location. As can be seen by examination of Figure 1, the toy waver is situated approximately 30 to 45 degrees to the participant's right, while a loudspeaker and display are 30 to 45 degrees to the participant's left.

The function of the toy waver is to maintain the attention of the test participant during the pre-stimulation phase of the test. Once the stimulation phase is initiated, the participant can be expected to shift his or her attention to the source of the stimulation which is the loudspeaker. Also not shown in Figure 1 is the test participant's parent, guardian or individual with whom the infant/toddler participant is familiar. The participant will be seated in the parent's lap so as to increase the likelihood that the participant will be comfortable and attentive during the test. The test set up is flexible so as to permit the participant to be seated independent of the parent, when this is practical.

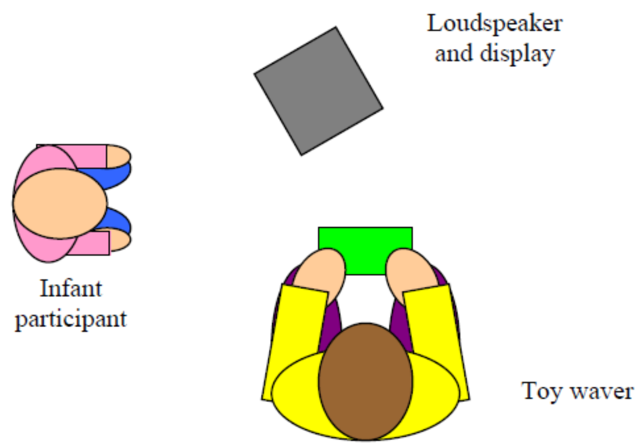

Figure 1. Plan of Test Layout

The process presently includes two test team members in addition to the test participant. The team is comprised of the test administrator and the toy waver. The test administrator's function is to manipulate the computer that generates the phonetic stream and records the responses of the participant. The phonetic stream is composed of a sequence of phonetic sounds that are presented as single syllables to the participant through a loudspeaker. In the sequence, a phonetic is repeated several times. At a given point in the sequence, the phonetic undergoes a subtle change, hereafter known as a phonetic change event. As an example, the phonetic "Lah" may be repeated until the test administrator initiates the phonetic change event. At this point, the phonetic "Bah" is repeated three times. This phonetic sequence may be better understood by examining Figure 2.

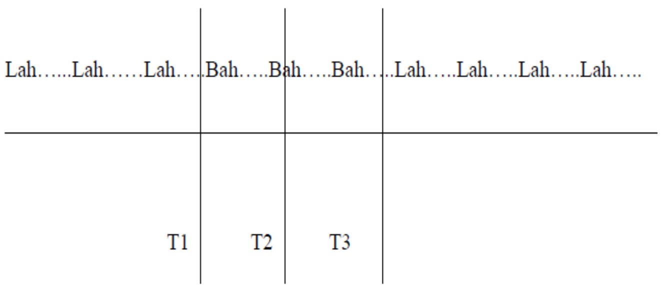

Figure 2. Test of the Phonetic Aspects of the Measurement

The times denoted as T1, T2 and T3 reference events which occur during the test. T1 denotes the change in phonetics used in the test. T2 occurs when the participant begins to move his or her head. T3 refers to the time when the participant's head is positioned at the stimulated orientation. The goal of the test is to detect the participant's reaction to the phonetic change event. Participant gaze reorientation, coincident with the phonetic change event toward the speaker from which the sounds are emanating, is recorded by the test administrator as a successful detection by the participant that a phonetic change event has occurred. The task of the administrator is a) to detect that a reorientation occurred and b) to ascertain that reorientation was a deliberate reaction of the participant to the phonetic change.

The toy waver's function is to attract and maintain the attention of the participant by use of actions and gestures with the toy, as shown in Figure 3. This orientation of the head and gaze toward the toy waver is referred to as the rest orientation or the rest position. In addition to these two members, the parent of the infant or toddler participant may be present to reduce tension in the participant. The test administrator is not visible to the toy waver, the participant or the parent of the participant. In this way, neither the toy waver nor the parent will know when the phonetic change event will take place. This is important since such pre-awareness might prompt the parent or toy waver to anticipate the movement of the participant's head with an inadvertent glance in the direction of the loudspeaker. Such actions have been demonstrated to be sensed by participants, a phenomenon referred to as gaze following [3]. Such gaze following would create an undesired control variable that would have a contaminating effect on the outcome of the test.

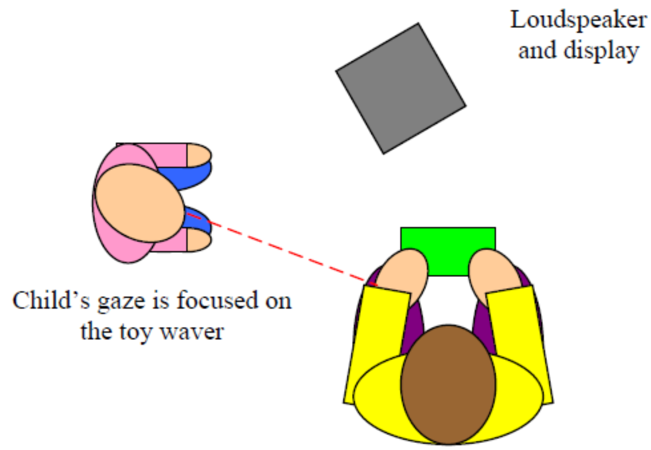

Figure 3. Participant's Attention Focused on the Toy

Test participants are males or females not younger than six months of age [4]. In the present test process, the infant or toddler participant is placed in the parent's lap in close proximity to the toy waver and to the loudspeaker and display. The toy waver maintains the concentration of the participant by showing the participant a toy, generally a stuffed animal [5]. Simultaneously, the computer software used to administer the test causes an audible repetition of a syllable which is common in the English language. An example is "Lah." The phonetic "Lah" is repeated at a rate of approximately once per two seconds. At an appropriate moment, the test administrator, through computer control, initiates the phonetic change event, at which point the repeated phonetic changes from "Lah" to another phonetic, "Bah," for example. The changed sound is repeated three times after which the software reverts back to the original phonetic.

With the change in the phonetic, the display that is on the same axis as the loudspeaker will present a picture of a stuffed animal. In this way the participant is presented with a pleasing stimulus as a reward for having noticed a change in the syllable that is being repeated. A head turn event is declared to have occurred when the participant deliberately shifts his or her attention to the speaker, as depicted in Figure 4, in reaction to the change in phonetic. This head and gaze orientation is referred to as the stimulated orientation or stimulated position.

Limitations of the Existing Test Regimen

In its present form, the test regimen makes use of a high level of human interface which limits accuracy and richness of the data that is captured from the test. Primarily, the occurrence of the head turn event is determined as the test administrator observes the movement of the participant’s head from the rest position to the stimulated position. It may be difficult to declare with certainty that the event actually occurred or that a movement of the head was a result of the participant’s detection of the phonetic change. In addition, no attempt is made in the present iteration to capture the time of the event. The importance of the detection and measurement of the latent times during the test shall be discussed later.

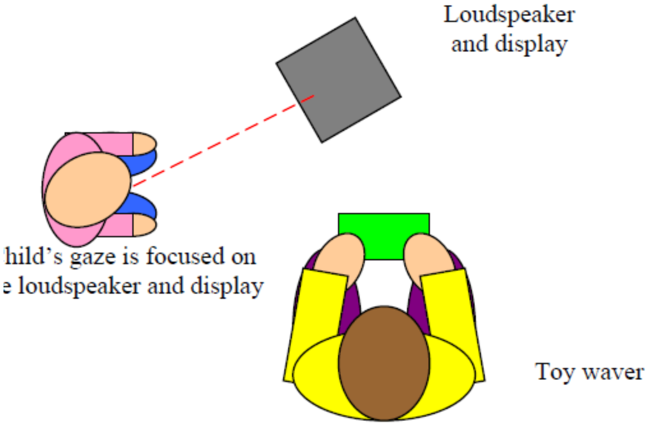

Figure 4: Participant’s Attention Focused on Loudspeaker

The Revision

The research team envisions the application of video imaging and motion analysis, referred to here as video-metrics, to the task of providing highly accurate positional data based on the video capture of the participant’s head as it moves from the rest position to the stimulated position. The field of video image analysis has assumed prominence in numerous disciplines such as security, sports and mechanical engineering. The research team has identified no project which intends to apply video-metric techniques to the HTT test process.

Lab Layout

In general terms, the technique involves making a video of the participant as the test proceeds using a digital video camera. The camera is interfaced to a computer which is running video image capture and analysis software. The image is sensed and quantified on a frame-by-frame basis. Instantaneous position of the participant’s head is analyzed to detect movement or changes in position over time. The revised test setup will be comprised of the components as described below. The participant station will continue to include a position for the toy waver, the participant and the loudspeaker and display. The test station is separated from the administrator’s position as in the present iteration. The layout of the lab is depicted in Figure 5.

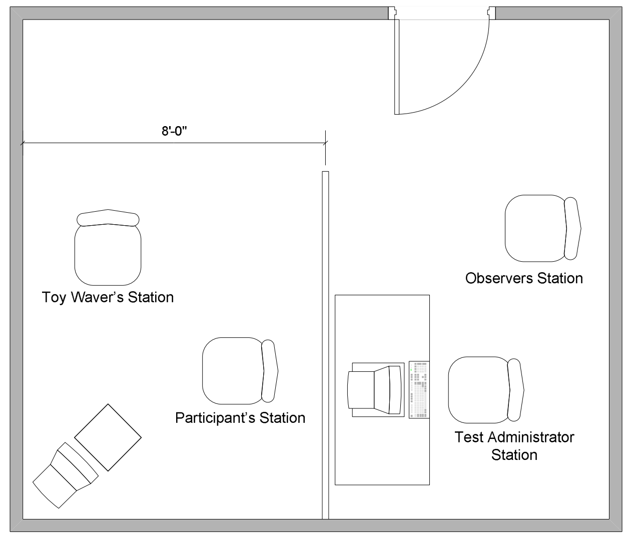

Figure 5: Plan View of Head-Turn Test Lab

The video camera will be situated in such a way as to be focused upon the participant and, more specifically, on the participant’s head. This camera will be interfaced with a standard desktop computer located at the test administrator’s position. In the present rendition, the single function of the computer is to control the transmission of the phonetic sounds to the loudspeaker. However, in the revision, the computer takes on most of the existing functions of the test administrator. In addition to controlling the audio, the computer will collect the video frames from the camera, process the video image frames, capture the metrics embodied in the frames, and store and display the data which emerge from the test. In addition, it is hoped that the test results will be presented in real time to the internet for viewing at remote locations.

Software Revisions

As previously stated, the existing test configuration makes use of computer support simply to manage the transmission of the audio stimuli. It is necessary to modify this software to provide additional capabilities as follows:

1. The revised software must embody a graphic user interface (GUI) that serves as a “dashboard” for the test administrator. The dashboard must give the test

administrator control over those functions necessary to conduct the test, store the test results, display the test results including an "instant replay" of the test video, mathematically analyze the test video from a time/event standpoint, and make the results available to external users including the Internet.

2. The revised software must be able to receive and process the video signal from the camera.
3. The revised software must be able to execute all functions of the existing software plus all of the functions of the test procedure as revised. This includes generation of the phonetic stimuli and timing management of the same, graphic presentation of the test participant focusing on movement of the head, control of the presentation of the toy on the display at the participant station, capture and analysis of the video image produced by the camera, storage of test results, replay of test results, generation and display of the timeline in a useable format, and statistical analysis of results of multiple iterations of the test.

The research team hopes to make use of off-the-shelf software for the video image analysis function. Another task which is part of the research is to tailor the available software for use in the HTT test environment. Additionally, the team will develop the GUI which will permit test administration and data analysis plus presentation of the test using the Internet. Generation, timing and control of the audio stimuli, as well as the image of the toy, should be a simple matter of configuration of the computer operating system. Researchers anticipate using Microsoft products for statistical analysis, storage and recall of test results.

Implications

As can be seen, the revision of the existing test proposed here provides users the ability to measure the speed with which the participant is able to process the perception of an audible stimulus and convert it to a completed head turn response. For the concept to be valid, accurate measurement of times and events must be an aspect of the regimen. This need for accuracy is predicated upon the premise that head position is an accurate indicator of focus of attention. This need not be assumed since Caron et al. [5] demonstrated the relationship in infants that have reached the age of 12 months. The proposed revision might facilitate measurements, suggested by Caron et al., of younger infants. Since a connection has been established between certain pathologies and impairment of the head turn function [6] the measurement of the times discussed above can be a useful diagnostic tool. Such conditions might be identified in participants. Therapies might be applied and then the effect of the therapies could be measured longitudinally by testing the head turn performance over time. The hereditary connection suggested by other researchers could he further explored [7].

There is cross-sectional value in determining the impact of various environmental factors upon such measurements. It may be inferred that the interval between T1 and T2 is useful in analyzing mental processing time. The interval between T2 and T3 could be an indicator of muscle control and optical/auditory performance. Indeed, there is potential value in the examination of head turn performance in older populations. The test setup permits the testing of such participants with little or no change in configuration and only minor changes in the test procedure. These are issues which will become more obvious as the revision is implemented; however, researchers anticipate the ability to measure the impact of substance abuse in participants as well as exposure to toxic or harmful substances. This has ramifications in the chemical and nuclear industries.

The technique could be applied to the Freiberg and Crassini [8] examination of infant sensitivity to sound power levels (SPL) with the addition of the ability to carefully measure absolute and relative sound power levels of the phonetic stimulus. The inquiry into an infant's ability to synchronize visual and audible stimuli by Hollich et al. [9] might be further facilitated by the enhancement to the Head Turn Technique.

In the initial proposal for enhancement [10], the research team established a standard for angular accuracy in the test results. The present proposal does not include a discussion of angular accuracy since part of the research will be to determine how angular accuracy may be established in the test procedure/setup. The benefits of attaching the angular accuracy discussed in the previous proposal will be discussed in future reports on the research as the project proceeds.

Conclusions

The disciplines involved in this project include video image capture and analysis, high- and low-level programming, audio signal technology, system analysis and project management. In addition, the application of test results is applicable to the social and medical fields. When applied to adults in the workplace as discussed above, HTT can serve as a tool to gather data for management or social programs. The test has broad applications in a wide range of disciplines. There are interesting implications for future research attaching to angular accuracy, which will find its genesis in the present project.

References

[1] Liu, H. M., Kuhl, P. K. & Tsoa, F. M. (2003). An Association Between Mother's Speech Clarity and Infant's Speech Discrimination Skills. *Developmental Science*, 6(3), F1–F10.

[2] Anderson, J. L., Morgan, J. L. & White, K. S. (2003). A Statistical Basis for Speech Sound Discrimination. *Language and Speech*, 46(2/3), 155–183.

[3] Brooks, R. & Meltzoff, A. N. (2002). The Importance of Eyes: How Infants Interpret Adult Looking Behavior. *Developmental Psychology*, 38(6), 958–966.

[4] Bortfield, A. A., Morgan, J., Golinkoff, R. M. & Rathburn, K. (2005). Mommy and Me. *Psychological Science*, 16(4), 298–304.

[5] Caron, A. J., Butler, S. & Brooks, R. (2002). Gaze Following at 12 and 14 Months: Do the Eyes Matter? *British Journal of Developmental Psychology*, 20(2), 225–240.

[6] Benasich, A. A. (2002). Impaired Processing of Brief, Rapidly Presented Auditory Cues in Infants With a Family History of Autoimmune Disorder. *Developmental Neuropsychology*, 22(1), 351–372.

[7] Choudhury, N., Leppanen, P. H. T., Leevers, H. J. & Benasich, A. A. (2007). Infant Information Processing and Family History of Specific Language Impairment: Converging Evidence for RAP Deficits from Two Paradigms. *Developmental Science*, 10(2), 213–236.

[8] Freiberg, K. & Crassini, B. (2002). Use of an Auditory Looming Task to Test Infants' Sensitivity to Sound Pressure Level as an Auditory Distance Cue. *British Journal of Developmental Psychology*, 19(1), 1–10.

[9] Hollich, G., Newman, R. & Jusczyk, P. W. (2005). Infants' Use of Synchronous Visual Information to Separate Streams of Speech. *Child Development*, 77 (3), 598–614.

[10] Hoy, B. A. & Hoy E. L. (2009). Enhancement to the Conditioned Head Turn Technique for Measuring Infant Response to Audio Stimulus. *International Journal of Modern Engineering*, 9(1), 43-51.

Biographies

BARRY HOY is an Associate Professor and Chair of Department of Human Resources and Health Care Management for Saint Leo University. He is currently the lead consultant of a grant from the Nuclear Regulatory Commission supporting continued research for the Head Turn Project. Dr. Hoy has more than 17 years of experience as an educator, corporate training director, and educational systems developer. He is also a retired naval officer. Barry Hoy may be reached at barry.hoy@saintleo.edu

ELEANOR HOY is the principal investigator for a Head Turn related grant funded by the Nuclear Regulatory Commission. Dr. Hoy is an assistant Professor in the School of Science, Engineering and Technology at Norfolk State University. She has more than 25 years experience as a technology educator at various institutions of higher learning. Eleanor Hoy may be reached at elhoy@nsu.edu