The Journal of International Advanced Otology

Int Adv Otol 2014; 10(2): 172-80 • DOI: 10.5152/iao.2014.118

Original Article

# Development of a Turkish Monosyllabic Word Recognition Test for Adults

**Serpil Mungan Durankaya, Bülent Şerbetçioğlu, Gökhan Dalkılıç, Selhan Gürkan, Günay Kırkım**

Department of Otorhinolaryngology-Head and Neck Surgery, Hearing, Balance and Speech Unit, Dokuz Eylül University Faculty of Medicine, İzmir, Turkey (SMD, BŞ, SG, GK)
Department of Computer Engineering, Dokuz Eylül University Faculty of Engineering, İzmir, Turkey (GD)

**OBJECTIVE:** The aim of this study was to develop a Turkish-specific speech recognition test, considering phonemic balance, homogeneity, and familiarity criteria.

**MATERIALS and METHODS:** The most frequently used Turkish monosyllabic words were selected from the corpus. Thirty-six young adults with normal hearing were divided into two groups and asked to listen to words from the word pool; the words were given twice at six different intensity levels. The least and most frequently known words were identified and eliminated in order to provide homogeneity. Three word lists, each composed of 50 phonemically balanced words, were developed to be used in the tests. These lists were divided into two according to the phonemic balance criteria.

**RESULTS:** No statistically significant difference was found among the word lists. Furthermore, the internal reliability of each list was analyzed using KR-20 and was found to be above 98% for all lists.

**CONCLUSION:** The lists derived from the Turkish language were ascertained to be appropriate for use. As a result of using the developed lists to test individuals with various auditory pathologies, it will be possible to assess the lists' capability to distinguish pathological cases according to the location of the pathology.

**KEY WORDS:** Monosyllabic word, speech recognition test, speech audiometry, word lists

## INTRODUCTION

Word recognition tests are used to evaluate listeners' capacity and speech in daily life. They are also used to aid in making differential diagnoses between peripheral and central auditory pathologies [1-4]. The use of appropriate speech materials is crucial to ensure the reliability and diagnostic sensitivity of word recognition tests [2, 5], which may be affected by a number of factors, such as the contents of the list, word selection homogeneity [6, 7], the accent of the speaker [5, 8-11], the method and level of presentation [12, 13], subject factors [13], and the type of recording [9]. Significant differences in speech audiometry results in hearing-impaired populations due to the speaker's gender have also been found in several studies, as well [5, 14].

Egan [15] specified the following criteria regarding the selection of word lists utilized in word recognition tests: monosyllabically structured words, equivalent average difficulty between lists, equivalent average difficulty within lists, equivalent phonetic composition between lists, a composition representative of the spoken language, and words that are commonly used. Frequently used words in a language are better recognized compared to those less frequently used [16]. The selected words to be included in the speech recognition tests should neither be very easy nor very difficult to distinguish [15, 17]. In order to assert the validity and reliability of a speech test, the principle of phonemic balance (PB) has been applied for a long time in the development of monosyllabic materials [16-18]. Other important principles in designing speech test materials are the familiarity of test items and the number of test items [2, 16, 19]. Typically, 50-items word lists are adopted in word recognition tests [2, 15, 18, 20]. The inclusion of 50 items increases the reliability of word recognition scores. Nevertheless, audiologists commonly use 25-items word lists to decrease test time and thereby reduce the fatigue effect for participants. Word familiarity is basically defined by a word's frequency of occurrence within the language, as reported in established word counts [8]. It is also imperative to utilize speech audiometry test materials in the native language of the subject being tested to obtain valid speech recognition test results [21]. For this purpose, speech tests have been developed for various languages, each to fit the requirements of measuring speech recognition based on the specific features of that language [22].

*This study was presented at the 5th National Audiology and Speech Disorders Congress, 23-26 September 2010, İzmir, Turkey.*
**Corresponding Address:**
Serpil Mungan Durankaya, Department of Otorhinolaryngology-Head and Neck Surgery, Hearing, Balance and Speech Unit Dokuz Eylül University Faculty of Medicine, İzmir, Turkey
Phone: +902324123296; E-mail: serpil.mungan@deu.edu.tr
**Submitted:** 27. 03. 2014 **Accepted:** 20.05.2014

Unfortunately, word recognition materials in the Turkish language have not been as well developed as those in English. Although Turkish is a language derived from the Ural-Altaic language family and is by some estimates the fifth-most widely spoken language in the world, to date, only a limited number of word recognition test materials have been available in Turkish. It is estimated that approximately 200 to 220 million people worldwide speak Turkish as their native or secondary language. The languages in the Ural-Altaic family have consonant/vowel harmony and an agglutinative structure, and there are no gender distinctions, such as feminine and masculine. Some inflectional morphemes can be used as derivational morphemes, and there are phonetic, morphological, and syntactic similarities among these languages [23].

The first word recognition test for Turkish in Turkey was developed more than 5 decades ago, and the speech recognition lists were not completely phonemically balanced, homogeneous, or familiar. The aim of the present study was to develop the open-set Turkish Monosyllabic Word Recognition Test (TMWRT) for adults.

## MATERIALS and METHODS

### Subjects

This study was started after receiving an approval report with the number 09/14/2009-138/2009 and dated June 18, 2009 of the Dokuz Eylül University Non-interventional Research Ethics Committee. Thirty-six subjects who were Turkish native speakers (18 men and 18 women) participated in an evaluation of the psychometric functions of the test material. The subjects ranged in age from 18 to 30 years (mean=22.8 years), and they had pure-tone air conduction thresholds of less than 10 dB HL at all octave frequencies from 250 to 8000 Hz and static acoustic admittance between 0.3 and 1.4 mmhos, with peak pressure between -100 and +50 daPa [24, 25]. For each subject, the ear with the better pure-tone average thresholds at 500 Hz, 1000 Hz, and 2000 Hz was selected as the test ear. Since inter-word and inter-talker variability is lower for male voices than female voices, the 345 most frequently used monosyllabic words were recorded by a male talker of age 55 using a standard dialect of the native language [6]. The selected 345 monosyllabic words were recorded by a 55-year-old male professional theater artist who has a standard dialect of Turkish language. Detailed information on the study was given to the subjects, and they asked to read and sign the consent form.

### Preparation of Test Materials

Monosyllabic Turkish words that have a consonant-vowel-consonant (CVC) structure were chosen as the word recognition test materials. Test materials were selected from Dokuz Eylül University, Computer Engineering Department's Corpus of Turkish, which comprises about 8 million monosyllabic words, referred to as the Turkish corpus [26]. Finally, the 500 most frequently counted monosyllabic words were selected as materials to design the word lists. Suggestiveness control of the chosen words was exercised through the *Turkish Language Association Dictionary* and *Written Turkish Word Frequency Dictionary* [27]. Words that were not found in either dictionary or were absurd, technical, or slang words were excluded. The residual 350 words were evaluated by 380 subjects with a variety of socio-economic and educational backgrounds, who were in the age range of 16 to 70. These subjects did not participate in this project other than through this task. These 350 monosyllabic words were then rated by 380 native judges on a scale of 1 to 3 based on how familiar a word would be to a Turkish speaker from Turkey (1=very familiar or frequently used, 2=infrequently used, 3=rarely used or not familiar). Finally, 5 out of 350 words that were considered to be culturally insensitive, unfamiliar, or inappropriate were excluded from the list; as a result, only 345 words were evaluated.

### Recordings

All recordings were performed in a large double-walled sound-treated booth in Dokuz Eylül University Faculty of Fine Arts, at a professional studio in Izmir, Turkey. A Neumann model U87 high-sensitivity microphone (Neumann U87; Neumann, Berlin, Germany) was positioned approximately 20 cm from the talker at 0° azimuth and was covered by a 10 cm windscreen. The microphone was connected to an AVALON AD2022 microphone preamp (Avalon AD2022; Avalon Design, Tustin, CA, USA), which was powered by an Avalon model B2T preamp power supply. The signal was digitized by a Digidesign Digi 002R 16-bit analog-to-digital converter (Digidesign; Avid Technology, Palo Alto, CA, USA) and subsequently stored on a hard drive for later editing. A 44100-Hz sampling rate with 16-bit quantization was used for all recordings, and every effort was made to utilize the full range of the 16-bit analog-to-digital converter.

During the recording session, the talker was asked to read each monosyllabic word at least four times with a slight pause between each production. Each monosyllabic word was recorded until two audiologists agreed that the vocal quality, accent, and pronunciation were satisfactory. The talker was asked to speak at a natural rate with a normal intonation pattern. After the rating process, the intensity of each word to be included in the listener evaluation was edited as a single utterance using Adobe Audition V:3 software (Adobe Audition; Adobe Systems Incorporated, San Jose, CA, USA) to yield the same average intensity as that of a 1000-Hz calibration tone.

### Procedures

Custom software (co-author; Dalkılıç, İzmir, Turkey) was used to control the randomization, presentation, and timing of the monosyllabic words to minimize practice effects for the subject. This software was developed by a co-author from the Department of the Computer Engineering for the purpose of presenting words in a random order. The speech stimulus was routed from a computer hard drive to the external input of an interacoustic model AC33 clinical audiometer (Interacoustic AC33; Interacoustics AS, Assens, Denmark). The stimuli were then routed from the audiometer to the participants via TDH-39 headphones. Audiological tests were performed in a double-walled sound-treated booth meeting the ANSI S3.1-1999 standards for maximum permissible ambient noise levels for the ears-not-covered condition using one-third octave-band measurements [28]. Prior to each subject's test session, the input to the audiometer was calibrated to 0 VU using the 1000-Hz calibration tone through customized software. In accordance with the American National Standards Institute's (ANSI) S3.6-2004 standards [29], the audiometer was also calibrated weekly during and at the conclusion of data collection. No changes in calibration were necessary throughout the course of data collection.

Each subject participated in two test sessions after passing a preliminary test. The 345 monosyllabic words were presented at six different intensity levels, beginning at 0 and ascending to 55 dB HL in 5-dB increments. Thirty-six subjects were divided into two groups (i.e., nine men and nine women in each group). The psychometric functions of the 345 most familiar monosyllabic words were established by presenting words at the levels of 0, 10, 20, 30, 40, and 50 dB HL to subjects in group 1 and at levels of 5, 15, 25, 35, 45, and 55 dB HL to subjects in group 2. At the end of the experiment, each subject listened to 2070 stimuli (i.e., the 345 most frequently presented monosyllabic words at six different levels).

Each subject was allowed to have several test periods during each test session. Prior to evaluation of the monosyllabic words, the following instructions were given:

"You will hear monosyllabic words, which may become louder or softer intensity. At the very soft levels, it may be difficult for you to hear words. Please listen carefully and repeat the words that you hear. If you are unsure of a word, you are encouraged to guess. If you have no guess, please say 'I don't know' and listen for the next word. Do you have any questions?"

The subjects were encouraged to guess and repeat what they heard or to say "I don't know" if they did not comprehend the stimuli at all. Their verbal responses were recorded with a Shure SM48 model dynamic microphone (Shure SM48; Shure Incorporated, Niles, IL, USA) on a computer for later assessment and analysis.

**Selection of Homogeneous Monosyllabic Words**

The psychometric functions of each of the 36 subjects and 345 most frequent monosyllabic words were fitted using third-degree polynomials ($y=a+bx+cx^2+dx^3$). A third-degree polynomial is an S-shaped curve described by four coefficients (a-d). The curve fitted to each monosyllabic word across the 36 subjects was then used to calculate the following four psychometric characteristics: (1) the threshold at 50% correct answers (i.e., the presentation level required for 50% word recognition); (2) the slope at 50% correct; (3) the slope between 20% and 80% correct; and (4) the intelligibility at 55 dB HL (i.e., word recognition score at the highest presentation level). The homogeneity of the word lists was achieved by examining the four psychometric characteristics of the 345 most frequent monosyllabic words. The following criteria were used: the monosyllabic words that yielded an intelligibility of at least 90% correct at 55 dB HL were included, and the monosyllabic words with intelligibility at 15 dB HL below 25% and over 95% were excluded. The 15 dB HL level was the one that yielded the most 50% correct responses for the selected monosyllabic words.

**Construction of the Word Lists**

To satisfy different clinical applications requiring 50- and 25-item word lists, six 25-item word lists were constructed, and these could be paired to form three phonemically balanced 50-item word lists. The design of the TMWRT word lists was aimed at not only ensuring phonemic balance but also at satisfying the application of 50- and 25- item word lists in different clinical situations. The phonemic balance of the 25-item word lists was achieved by equally dividing the desired numbers of initials in the 50-item word lists into two groups, called the half-A and half-B lists.

During the homogeneity stage, 61 words with unequal daily usage and dissimilar intensity-dependent repeatabilities were eliminated; 284 words proceeded to the list construction stage. In order to ensure the content validity of the word lists, phonemic balance was the preferred linguistic measure, as it is regarded as a both significant and frequently used method.

Considering the distribution of phonemes whose frequencies at the beginning, middle, and end of the word were predetermined in the corpus, three different phonemically balanced lists composed of 50 words were constructed. These lists were divided into two in a phonemically balanced manner, and as a result, six lists, each composed of 25 words, were obtained.

**Statistical Analyses**

16.0 (Statistical Package for the Social Sciences (SPSS) for Windows 16.0, SPSS Inc. 2007, Microsoft; Armonk, New York, USA) program was employed for statistical analysis of the study data. The analyses were carried out in a confidence interval of 95%, and $p<0.05$ was regarded as statistically significant. The dependent variable, defined as correct word recognition, was saved in binary format (correct versus incorrect). The independent variables were different intensity levels and different word lists. The raw data were used in logistic regression analysis. The raw scores of each word in the four lists were used in logistic regression analysis to calculate the logistic regression parameters for each list and half list. After the lists and half lists were compiled, a logistic regression yielding a chi-square ($\chi^2$) statistic was performed in order to determine whether there were any significant differences among the lists or half lists. Kuder-Richardson Formula 20 (KR-20) analyses were used for within-list equivalency. It is analogous to Cronbach's α, except that Cronbach's α is used for non-dichotomous measures.

**RESULTS**

In total, 36 individuals participated in this study. Pure tone threshold average (PTA) indicated that all participants had satisfied the criteria of having a PTA level less than 10 dB HL, with a mean PTA of 2.31±2.70 dB HL. Twenty-five participants were tested on the right ear, and eleven were tested on the left ear. The average pure tone threshold for ears, the average age, and a detailed description of the participants are presented in Table 1.

Following the raw data collection, logistic regression was used to obtain the regression slope and intercept for each of the 345 monosyllabic words. These values were then inserted into a modified logistic regression equation that was designed to calculate the percent correct at each intensity level. The original logistic regression equation is as follows:

$$log \frac{p}{1-p} = a + b \times i \quad (1)$$

In Equation 1, *p* is the proportion corrects at any given intensity level, *a* is the regression intercept, *b* is the regression slope, and *i* is the presentation level in dB HL. Percent correct values were then used to construct psychometric functions. When Equation 1 is solved for *p* and multiplied by 100, Equation 2 is obtained, where *p* is percent correct recognition:

$$p = (1 - \frac{\exp{(a+b \times i)}}{1+\exp{(a+b \times i)}}) \times 100 \quad (2)$$

**Table 1.** Participants with normal hearing: age, gender, test ear and average values for pure tone threshold and static compliance

| Group | Gender | | Age (year) | Test ear | | PTA[1] (dB HL) | PTA[2] (dB HL) | Static comp. ($mmH_2O$) |
|---|---|---|---|---|---|---|---|---|
| | Men | Female | | Right | Left | | | |
| 1 | 9 | 9 | 22.28±3.84 | 13 | 5 | 1.72±3.12 | 0.50±3.03 | 0.5±0.2 |
| 2 | 9 | 9 | 23.22±5.08 | 12 | 6 | 2.89±2.14 | 2.72±2.21 | 0.6±0.2 |

$PTA_1$: average of pure tone thresholds at frequencies of 0.5, 1, and 2 kHz
$PTA_2$: average of pure tone thresholds at frequencies of, 1, 2, and 4 kHz
Static comp: static compliance

**Table 2.** Mean values of the psychometric characteristics for the 345 most frequently occurring monosyllabic words, 284 homogeneous monosyllabic words, 50-item word lists and 25-item word lists

| Psychometric characteristic | 50% Threshold (dB HL) | 50% Slope (%/dB) | 20-80% Slope (%/dB) | Percentage of intelligibility at 55 dB HL |
|---|---|---|---|---|
| 345 monosyllabic words | 14.2 | 5.5 | 4.7 | 98.84 |
| 284 monosyllabic words | 13.4 | 6.1 | 5.3 | 98.94 |
| 50-item word lists | 13.5 | 6.2 | 5.4 | 98.67 |
| 25-item word lists | 13.5 | 6.2 | 5.4 | 98.67 |

**Table 3.** Mean performance of Turkish monosyllabic 50-item and 25-item word lists

| List | a[a] | b[b] | Slope at 50%[c] | Slope from 20 to 80%[d] | Threshold dB HL[e] |
|---|---|---|---|---|---|
| A | 3.4149 | -0.2500 | 6.3 | 5.4 | 13.7 |
| B | 3.3277 | -0.2471 | 6.2 | 5.4 | 13.5 |
| C | 3.3611 | -0.2484 | 6.2 | 5.4 | 13.5 |
| Mean | 3.3679 | -0.2485 | 6.2 | 5.4 | 13.5 |
| A1 | 3.3337 | -0.2516 | 6.3 | 5.4 | 13.3 |
| A2 | 3.4675 | -0.2492 | 6.2 | 5.4 | 13.9 |
| B1 | 3.3883 | -0.2547 | 6.4 | 5.5 | 13.3 |
| B2 | 3.2750 | -0.2403 | 6.1 | 5.2 | 13.6 |
| C1 | 3.3080 | -0.2472 | 6.2 | 5.4 | 13.4 |
| C2 | 3.3778 | -0.2496 | 6.2 | 5.4 | 13.5 |
| Mean | 3.3584 | -0.2488 | 6.2 | 5.4 | 13.5 |

a[a]: regression intercept; b[b]:regression slope. Slope at 50%; c: Psychometric function slope (%/dB) at 50% was calculated from 49.999 to 50.001%. Slope from 20 to 80%; d: Psychometric function slope (%/dB) from 20-80%.
Threshold dB HLe: Intensity required for 50% intelligibility

By inserting the regression slope, regression intercepts, and presentation level into Equation 2, it is possible to predict the percentage correct at any specified intensity level. Percentage of correct word recognition was predicted for each of the monosyllabic word full lists and half lists for a range of 0 to 55 dB HL in 2-dB increments. Psychometric functions were then produced using the predicted percentages.

The third-degree polynomials analysis was carried out with 345 words, after the elimination of 61 words that did not satisfy the homogeneity features, while 50% mean intensity, 50% slope, and 20%-80% slope were determined using 284 words. The mean values of the psychometric characteristics for the 345 most frequently occurring monosyllabic words, 284 homogeneous monosyllabic words, 25-item word lists, and 50-item word lists are given in Table 2.

From the three 50-item word lists, a 6.2%/dB, 50% slope, 5.4%/dB 20-80% slope, and 13.5-dB HL 50% mean intensity were obtained, and from the six 25-item word lists, a 6.2%/dB, 50% slope, 5.4%/dB 20-80% slope, and 13.5-dB HL 50% mean intensity were obtained. The intensity adjustments made to each word in the three 50-item word lists and six 25-item word lists are presented in Table 3.

As a part of the logistic regression analysis of the full lists and half lists, a two-way chi-square ($\chi^2$) analysis (intensity and list as independent variables with response as the dependent variable) was used to evaluate the inter-list equivalence among the six 25-item word lists and three 50-item word lists-that is, to determine whether the percent correct values differed significantly at the 12 different presentation levels. No significant differences were found among the word lists [$\chi^2(5) = 0.3433$ and p=0.634 for the six 25-item word lists and $\chi^2(2) = 0.386$ and p=0.824 for the three 50-item word lists]. The $\chi^2$ tests indicated that the six 25-item word lists and three 50-item word lists exhibited inter-list equivalence. At this stage, within-list consistency was also checked using Kuder-Richardson 20 (KR-20) consistency analysis. Results of the analysis were as follows: 0.982 for list A, 0.982 for list B, 0.983 for list C, 0.982 for list A1, 0.983 for list A2, 0.982 for list B1, 0.983 for list B2, 0.982 for list C1, and 0.983 for list C2. A list analysis yielded high consistency figures.

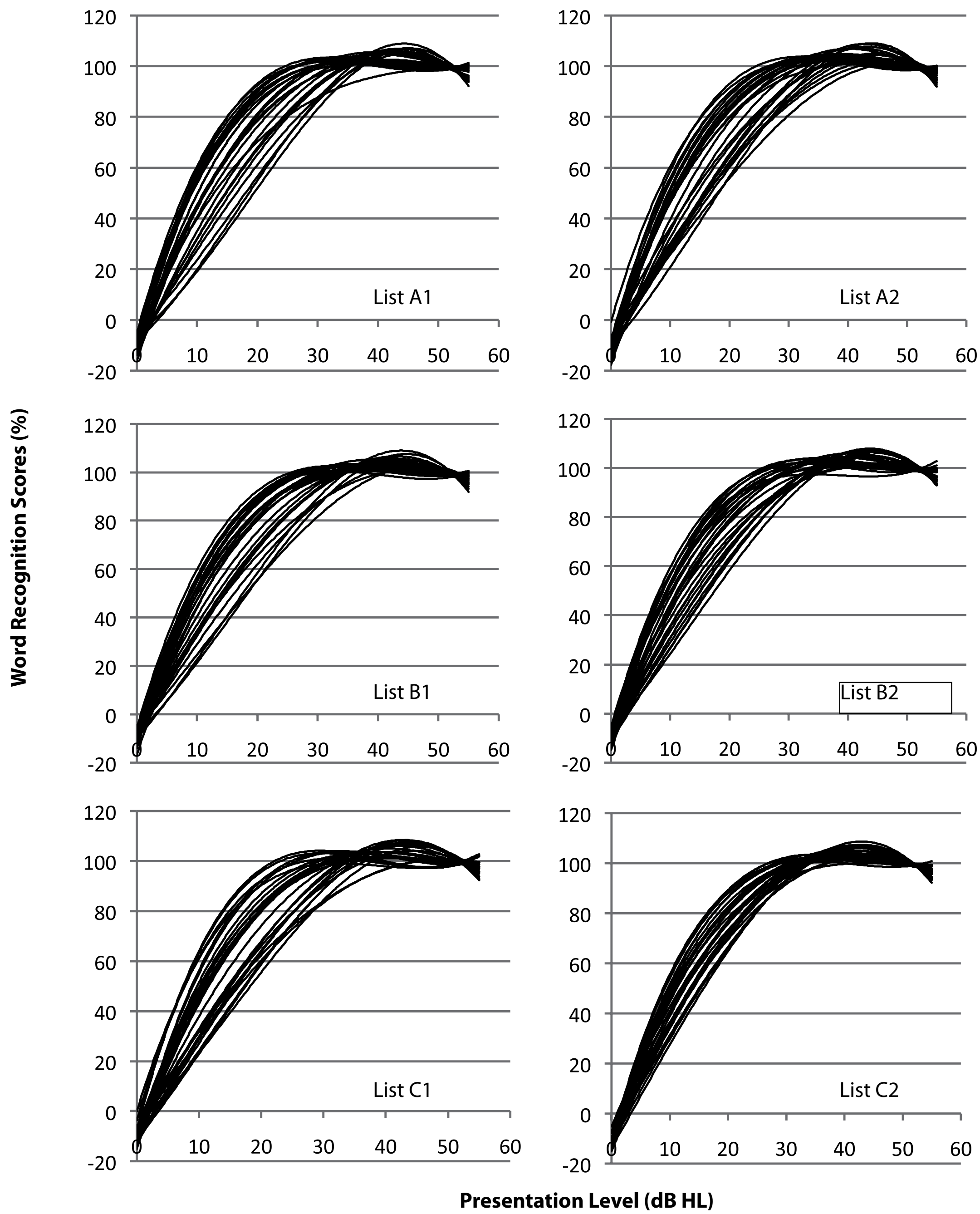


**Figure 1.** Psychometric functions for individual monosyllabic words obtained for six 25-item word lists

Furthermore, the internal reliability of each list was analyzed using KR-20 and was found to be above 98% for all lists. The psychometric functions of the six 25-item word lists are shown in Figures 1 and 2. The six 25-item word lists were called A1, A2, B1, B2, C1, and C2. Table 4 presents the words used in each half list in Turkish with their phonetic transcription, as well as the meanings of the words in English.

Words were distributed evenly within the lists, depending on their phonemic balance and frequency of occurrence within the lists. The frequencies of occurrence of the phonemes are presented in Tables 5 and 6 [26].

## DISCUSSION

The main purpose of the current study was to create a set of half and full word lists with homogeneous, psychometric equivalence, familiarity, and phonemic balance criteria for the Turkish Monosyllabic Word Recognition Test (TMWRT). Monosyllabic words are often used in speech recognition tests performed in clinical studies. It can be said that the knowledge associated with speech at the monosyllabic word level is quite low, and reduction of excessive knowledge causes auditory factors to come to the forefront in the recognition of words. Among monosyllabic words, words with a consonant-vowel-consonant (CVC) structure are most often used in Turkish [26]. The corpus studies show

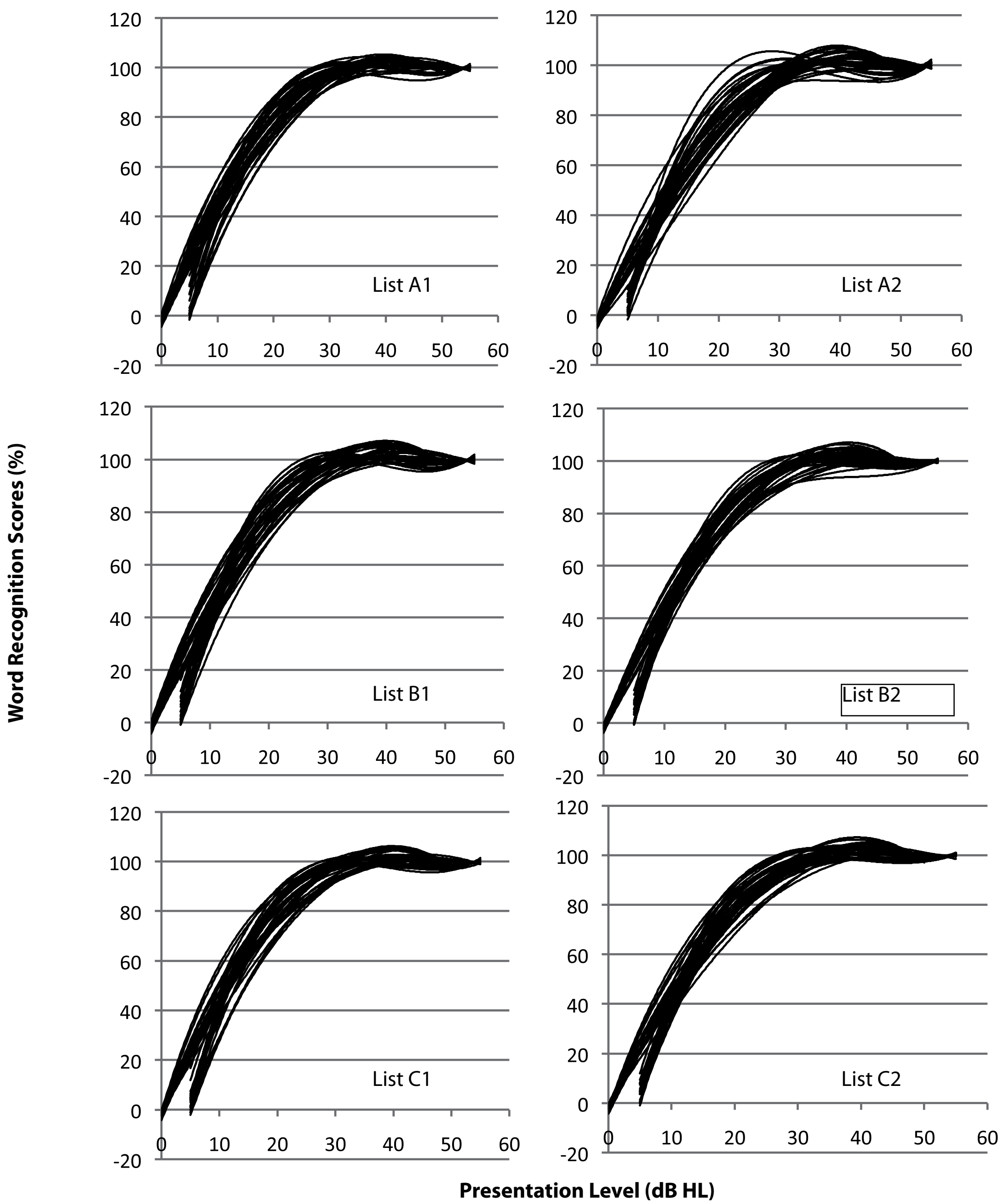


**Figure 2.** Psychometric functions for individual subjects obtained for six 25-item word lists

that the MLU (mean length of utterance) in Turkish is 6.34, and an analysis of CV order and word length showed that the orders of CVCVC, CVC, CV, CVCV, and CVCCV are the most common word structures [26, 30]. The frequency of CVC-ordered monosyllabic words was 7.25%, while the frequency of CV-ordered monosyllabic words was 6.92%, and the frequency of VC-ordered monosyllabic words was 1.08% in Turkish [26, 30]. These results show that the CVC order is the most common order among the monosyllabic words in Turkish. For this reason, only words with a CVC structure were employed in this study.

While Cevansir [31] and Cura [32] developed the word lists, Kılıncarslan [33] standardized the PB-300 lists. However, all of these researchers tried to employ monosyllabic words in general. In addition, two-, three-, and four-phoneme words have been used that did not have a consonant-vowel-consonant structure, and inter-list and within-list balance could not be ensured in such cases. Aksit [34] has constructed lists in the consonant-vowel-consonant format, using words in the lists standardized in 1986, as well as 40 words he obtained as a result of a dictionary scan.

Table 2 and Figures 1 and 2 indicate that the six 25-item word lists and the three 50-item word lists exhibit performance homogeneity in terms of inter-item variability, inter-subject variability, and four assessed psychometric characteristics. A $\chi^2$ test was used to evaluate the inter-list equivalence among the six 25-item word lists and the

**Table 4.** 25-item word lists and meaning of the words

| No | List A[1] | Meaning of words | List A[2] | Meaning of word | List B[1] | Meaning of word | List B[2] | Meaning of word | List C[1] | Meaning of words | List C[2] | Meaning of word |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Cep | Pocket | Çek | Check | Can | Life | Çay | Tea | Cin | Genie | Çal | Play |
| 2 | Rey | Vote | Pis | Filthy | Mor | Purple | Ruh | Soul | Nem | Humidity | Ray | Rail |
| 3 | Zor | Hard | Yay | Bow | Yar | Cliff | Zil | Bell | Yok | Absent | Zar | Membrane |
| 4 | Çam | Pine | Dev | Giant | Çık | Exited | Dar | Norrow | Çat | Snap | Dem | Well-steeped tea |
| 5 | Giy | Wear | Göç | Migration | Fes | Fez | Geç | Late | Fay | Fault | Gir | Enter |
| 6 | Buz | Ice | Bit | Bit | Bez | Cloth | Boy | Length | Bay | Mister | Bar | Bar |
| 7 | Bel | Waist | Biz | We | Bey | Esquire | Ben | I | Bol | Plenty | Boz | Gray |
| 8 | His | Feel | Kul | Servant | Pes | Low | Kül | Cinder | Hem | Both | Kes | Cut |
| 9 | Yer | Place | Yün | Wool | Yat | Yacht | Yem | Bait | Yak | Fire | Yel | Wind |
| 10 | Düş | Dream | Bor | Boron | Bil | Know | Bir | One | Bin | Thousand | Bul | Find |
| 11 | Dağ | Mountain | Din | Religion | Diş | Tooth | Diz | Knee | Del | Drill | Dök | Pour |
| 12 | Dik | Vertical | Far | Lamp | Dam | Roof | Gez | Visit | Dur | Stop | Güç | Power |
| 13 | Hat | Line | Her | Each | Gök | Sky | Kem | Evil | Gel | Come | Hız | Speed |
| 14 | Kum | Waist | Mum | Candle | Kor | Core | Nur | Light | Küt | Bump | Mit | Myth |
| 15 | Yön | Way | Tak | Attach | Vur | Hit | Yap | Do | Vah | Woe | Yan | Side |
| 16 | Kas | Muscle | Küp | Cube | Hür | Free | Kat | Floor | Kar | Snow | Kir | Dirt |
| 17 | Tan | Dawn | Şah | Shah | Şen | Merry | Tür | Kind | Ter | Sweat | Şov | Show |
| 18 | Bal | Honey | Doz | Dose | Dal | Branch | Dön | Turn | Dış | External | Dün | Yesterday |
| 19 | Set | Set | Sel | Flood | Sat | Buy | Sık | Frequent | Sür | Drive | Sun | Offer |
| 20 | Ver | Give | Tel | Wire | Tur | Tour | Tek | Alone | Ten | Skin | Tas | Bowl |
| 21 | Gün | Day | Baş | Head | Bağ | Connect | Boş | Empty | Baz | Base | Beş | Five |
| 22 | Mal | Ware | Nar | Pomegranate | Kin | Hatred | Has | Special | Mis | Fragrance | Pir | Patriarch |
| 23 | Kır | Prairie | Kan | Blood | Kap | Grab | Mil | Mile | Kur | Exchange | Kek | Cake |
| 24 | Sor | Ask | Sır | Secret | Sev | Love | Sar | Winding | Sap | Handle | Sağ | Right |
| 25 | Bak | Look | Ger | Stretch | Gül | Rose | Han | Hostelry | Gör | View | Hep | Always |

**Table 5.** Frequency of the occurence of vowels in the middle of words in the 50 item lists

| Vowels | Frequency of occurence in the middle of the word |
|---|---|
| A | 15/50 |
| E | 12/50 |
| I | 2/50 |
| İ | 7/50 |
| O | 4/50 |
| Ö | 2/50 |
| U | 4/50 |
| Ü | 4/50 |

three 50-item word lists-that is, to determine whether the values of percent correct recognition are distributed identically at the 12 different presentation levels (with list and presentation level as independent variables and percent correct recognition as the dependent variable). No significant differences were found among the word lists. The $\chi^2$ tests indicated that the six 25-item word lists and three 50-item word lists exhibited inter-list equivalence.

Similar slopes were obtained among the lists, although the determined slope was different from that of some of the languages in the literature. This difference may have been caused by the specific linguistic characteristics of Turkish. Turkish is an agglutinative language. Adding suffixes to fixed word roots in agglutinative languages can create word derivations and inflections. Monosyllabic words are the simplest word forms. Suffixes are used in Turkish in order to make new words [23, 35]. There can be trend differences between tests, even in the same language, based on both the participants and test standards.

The slope's lack of variability among 284 words and from list to list proves that homogeneity and inter-list balance were achieved in this study. The results of the speech recognition tests previously developed in Turkish were given in terms of percentages, and the slopes were not stated. For this reason, no clear comparison can be made with the previous Turkish studies. However, in similar papers published regarding other languages, comparisons were made on the basis of the slopes for different languages. In the foreign literature, slopes can be compared.

The slopes ranged from 20% to 80% for the 25-item monosyllabic word list and from 5.2%/dB to 5.7%/dB (m=5.4%/dB) and ranged from 5.3%/dB to 5.4%/dB (m=5.4%/dB) for the 50-item monosyllabic word list. The means that the slopes from the 20% to 80% range for the Turkish monosyllabic psychometric functions are close to those of word recognition test materials that have been reported in other

**Table 6.** Distribution of consonants and their frequencies at the beginning and end of the 50-item lists

| Consonants | Frequency of consonants at the beginning of the words | Frequency of consonants at the end of the words |
|---|---|---|
| B | 8/50 | 0/50 |
| C | 1/50 | 0/50 |
| Ç | 2/50 | 1/50 |
| D | 6/50 | 0/50 |
| F | 1/50 | 0/50 |
| G | 4/50 | 0/50 |
| Ğ | 0/50 | 1/50 |
| H | 3/50 | 1/50 |
| J | 0/50 | 0/50 |
| K | 6/50 | 4/50 |
| L | 0/50 | 6/50 |
| M | 2/50 | 3/50 |
| N | 1/50 | 6/50 |
| P | 1/50 | 2/50 |
| R | 1/50 | 11/50 |
| S | 4/50 | 3/50 |
| Ş | 1/50 | 2/50 |
| T | 3/50 | 3/50 |
| V | 1/50 | 1/50 |
| Y | 4/50 | 3/50 |
| Z | 1/50 | 2/50 |

**Table 7.** The mean slopes for English and other languages (%/dB)

| Languages | Tests and Authors | Slopes (%/dB) |
|---|---|---|
| Mandarin (Male Speaker) | Tsai et al. [2] | 4.1 |
| Spanish | Weisleder and Hodgson [9] | 4.3 |
| Mandarin | Han et al. [16] | 4.1 |
| English | NU-6, Beattie et al. [36] | 4.2 |
| English | CIDW-22, Beattie et al. [36] | 4.6 |
| English | NU-6, Wilson and Oyler [37] | 4.4 |
| English | CIDW-22 Wilson and Oyler [37] | 4.8 |
| English | NU-6, Tillman and Carhart [38] | 5.6 |
| English | NU-6,Wilson et al. [39] | 3.6 |
| Korean (Male speaker) | Harris et al. [40] | 5.0 |
| Korean (Female speaker) | Harris et al. [40] | 5.1 |
| Arabic | Alusi et al. [41] | 5.0 |
| Polish (Male speaker) | Harris et al. [42] | 5.8 |
| Polish (Female speaker) | Harris et al. [42] | 5.9 |
| Russian (Male speaker) | Harris et al. [43] | 5.8 |
| Russian (Female speaker) | Harris et al. [43] | 5.6 |
| Turkish (Male speaker) | Current study | 5.4 |

languages. The mean slope of monosyllabic words in English [36] was estimated as 4.6%/dB for CID W-22 and 4.2%/dB for the Northwestern University Auditory Test No. 6 word lists, respectively; Wilson and Oyler [37] stated that they were 4.8%/dB and 4.4%/dB. A comparison of the mean slopes for English and other languages is presented in Table 7 [2, 9, 16, 36-43].

Using homogeneous words in the development of word lists reduces both inter-item and inter-participant variability and increases inter-list equivalence [44]. Hood and Poole [17] have stated that difficult- and easy-to-distinguish words change the speech recognition percentage, so that word lists composed of words with average difficulty should be used. To ensure homogeneity, this variability has been reduced in our study.

As regards the limitations of the study, the total number of participants of the study was 30 normal hearing subjects. For this kind of preliminary study, 30 participants was considered sufficient; however, in order to validate, this word recognition list has to be tested in not only normally hearing subjects but also patients with different types of hearing losses. There is a need for further investigation and development of additional Turkish word recognition test materials. This article details an initial study on the creation of word recognition test materials for native speakers of Turkish. There continues to be a need for additional research on word recognition tests and the area of speech audiometry.

In conclusion, to develop a speech recognition test in this study, criteria, such as familiarity, homogeneity, phonemic balance, and inter-list balance, were taken into consideration in the selection of the words to be used. Implementing a speech recognition test by employing this subsequent voice recording and the software prepared at every clinic using these materials was appropriate with respect to ensuring test-retest standardization and avoiding test executor-dependent variables.

The results of the research related to Turkish speech recognition tests in Turkey are comparable with studies of other languages in the literature. Thus, conducting more studies on Turkish speech tests and making slope comparisons will be beneficial in terms of improving the Turkish speech recognition test.

We also realize that word recognition test materials are usually made equivalent based on adults with normal hearing. As a result of using the developed lists to test individuals with various auditory pathologies, however, it will be possible to assess the lists' capability to distinguish pathological cases according to the location of the pathology. Furthermore, it will be possible to realistically examine the daily verbal communicative problems of individuals with hearing loss.

**Ethics Committee Approval:** Ethics committee approval was received for this study from the ethics committee of Dokuz Eylül University's Non-interventional Research Ethics Committee committee (18.06.2009-09/14/2009-138).

**Informed Consent:** Written informed consent was obtained from patients who participated in this study.

**Peer-review:** Externally peer-reviewed.

**Author Contributions:** Concept - B.S., G.K., S.M.D., G.D., S.G.; Design - B.S., G.K., S.M.D., G.D., S.G.; Supervision - B.S., G.K.; Funding - S.M.D., B.S., S.G.; Materials - S.M.D., G.D., S.G.; Data Collection and/or Processing - S.M.D., G.D., S.G.; Analysis and/or Interpretation - G.D., S.M.D.;

Literature Review - S.M.D., S.G.; Writing - S.M.D., S.G., G.D., B.S., G.K.; Critical Review - B.S., G.K., G.D.

**Conflict of Interest:** No conflict of interest was declared by the authors.

**Financial Disclosure:** The authors declared that this study has received no financial support.

## REFERENCES

1. Egan JJ. Basic aspects of speech audiometry. Ear Nose Throat J 1979; 58: 190-3.
2. Tsai KS, Tseng LH, Wu CJ, Young ST. Development of a Mandarin monosyllable recognition test. Ear Hear 2009; 30: 90-9. [CrossRef]
3. Ma X, McPherson, B., Ma, L. Chinese speech audiometry material: Past, present, future. Hearing, Balance and Communication 2013; 11: 52-63. [CrossRef]
4. Ji F, Xi X, Chen AT, Ying J, Wang QJ, Yang SM. Development of a Mandarin monosyllable test material with homogenous items (I): Homogeneity selection. Acta Oto-Laryngol 2011; 131: 962-9. [CrossRef]
5. Nissen SL, Harris RW, Slade KB. Development of speech reception threshold materials for speakers of Taiwan Mandarin. Int J Audiol 2007; 46: 449-58. [CrossRef]
6. Cambron NK, Wilson RH, Shanks JE. Spondaic word detection and recognition functions for female and male speakers. Ear Hear 1991; 12: 64-70. [CrossRef]
7. Wilson RH, Carter AS. Relation between slopes of word recognition psychometric functions and homogeneity of the stimulus materials. J Am Acad Audiol 2001; 12: 7-14.
8. Penrod, J.P. Speech threshold and word recognition/discrimination testing. In: Katz J (Ed.) Handbook of Clinical Audiology, Fourth Edition, USA, Williams & Wilkins, 1994; 147-64.
9. Weisleder P, Hodgson WR. Evaluation of four Spanish word-recognition-ability lists. Ear Hear 1989; 10: 387-92. [CrossRef]
10. Carhart, R. Problems in the measurement of speech discrimination. Arch Otolaryngol 1965; 82: 253-60. [CrossRef]
11. Luce PA. A computational analysis of uniqueness points in auditory word recognition. Percept Psychophys 1986; 39: 155-8. [CrossRef]
12. Kent RD, Wiley TL, Strennen ML. Consonant discrimination as a function of presentation level. Audiology 1979; 18: 212-24. [CrossRef]
13. Boothroyd A. Developments in speech audiometry. Sound 1968; 2: 3-10.
14. Mackersie CL. Tests of speech perception abilities. Otolaryngol Head Neck Surg 2002; 10: 392-7.
15. Egan JP. Articulation testing methods. Laryngoscope 1948; 58: 955-91. [CrossRef]
16. Han D, Wang S, Zhang H, Chen J, Jiang W, Mannell R, et al. Development of Mandarin monosyllabic speech test in China. Int J Audiol 2009; 48: 300-11. [CrossRef]
17. Hood JD, Poole JP. Influence of the speaker and other factors affecting speech intelligibility. Audiology 1980; 9: 434-55. [CrossRef]
18. Lehiste I, Peterson G. Linguistic considerations and intelligibility. J AcoustSoc Am 1959; 31: 280-6. [CrossRef]
19. Shi LF, Sánchez D. The role of word familiarity in Spanish/English bilingual word recognition. Int J Audiol 2011; 50: 66-76. [CrossRef]
20. Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson RW. Development of materials for speech audiometry. J Speech Hear Disord 1952; 17: 321-37.
21. Ramkissoon I. Speech recognition thresholds for multilingual populations. Communication Disorders Quarterly 2001; 22: 158-62. [CrossRef]
22. Abdulhaq NMA. Speech Perception Test for Jordanian Arabic Speaking Children. University of Florida. Degree of Doctor of Philosophy 2006.
23. Özyetkin A.M. "Tarihten Bugüne Türk Dili Alanı," (Conference) Chinese Academy of Social Science, Sino-Foreign Relationship Department of Institute of History, Beijing/China, 2006.
24. American Speech-Language-Hearing Association. Guidelines for screening for hearing impairments and middle-ear disorders. ASHA 1990; 32: 17-24.
25. Roup CM, Wiley TL, Safady SH, Stoppenbach DT. Tympanometric screening norms for adults. Am J Audiol 1998; 7: 55-60. [CrossRef]
26. Orucu, F. Turkish Language Characteristics and Author Identification. Izmir, Dokuz Eylul University, Master of Science, 2009.
27. Göz İ. Yazılı Türkçenin Kelime Sıklığı Sözlüğü. Ankara, Ankara Türk Dil Kurumu Yayınları, 2003.
28. American National Standards Institute. Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms. New York, ANSI, 1999; S3.1
29. American National Standards Institute. Specification for Audiometers. ANSI S3.6-2004. New York, ANSI, 2004.
30. Dalkılıc G, Çebi Y. Türkçe külliyat oluşturulması ve Türkçe metinlerde kullanılan kelimelerin uzunluk dağılımlarının belirlenmesi. DEÜ Mühendislik Fakültesi Fen ve Mühendislik Dergisi, 2003; 5: 1-7.
31. Cevansir, B. Konuşma Odiometrisi Kelime ve Sayı Testleri. İstanbul, Istanbul University. Assoc. Prof. Thesis, 1965.
32. Cura, O, Günhan, Ö, Palandöken, M. Yeni Türkçe koklear kelime listelerinin takdimi (Yeni istatistiksel verilere dayanılarak, Türk Dili fonemi geçerliliğine göre eski listeler üzerinde yapılan değişiklikler). İzmir Devlet Hastanesi Mecmuası, 1976; 14: 1-49.
33. Kılıncarslan A. Türk Dili için Geliştirilmiş Fonetik Dengeli Tek Heceli Kelime Listelerinin Standardizasyonu. Ankara, Hacettepe University,Master of Science, 1986.
34. Aksit M. Konuşmayı Ayırt Etme Testi İçin İzofonik Tek Heceli Kelime Listelerinin Oluşturulması. Istanbul, Marmara University, Master of Science. Unpublished data, 1994.
35. Aksan D. Türkiye Türkçesinin Dünü, Bugünü, Yarını, Bilgi, Ankara, 2000; 1-68.
36. Beattie RC, Edgerton BJ, Svihovec DV. A comparison of the audiotec of St. Louis cassette recordings of NU-6 and CID W-22 on a normal-hearing population. J Speech Lang Hear Disord 1977; 42: 60-4.
37. Wilson RH, Oyler A. Psychometric functions for the CID W-22 and NU Auditory Test No. 6. Materials spoken by the same speaker. Ear Hear 1997; 18: 430-3. [CrossRef]
38. Tillman TW, Carhart R. An expanded test for speech discrimination utilizing CNC monosyllabic words. Northwestern University Auditory Test No. 6. SAM-TR-66-55. Tech Rep SAM-TR. 1966: 1-12.
39. Wilson RH, Coley KE, Haenel JL, Browning KM. Northwestern University auditory test no. 6: normative and comparative intelligibility functions. J Am Audiol Soc 1976; 1: 221-8. [CrossRef]
40. Harris RW, Nielson WS, McPherson DL, Skarzynski H, Eggett DL. Psychometrically equivalent Polish monosyllabic word recognition materials spoken by male and female talkers. Audiofonologia, 2004; 25, 16-31.
41. Alusi HA, Hinchcliffe R, Ingham B, Knight JJ, North C. Arabic speech audiometry. Audiology 1974; 13: 212-30. [CrossRef]
42. Harris RW, Nielson WS, McPherson DL, Skarzynski H, Eggett DL. Psychometrically equivalent Polish monosyllabic word recognition materials spoken by male and female talkers. Audiofonologia, 2004; 25, 16-31.
43. Harris RW, Nissen SL, Pola MG, McPherson DL, Tavartkiladze GA, Eggett DL. Psychometrically equivalent Russian speech audiometry materials by male and female talkers. International Journal of Audiology 2007; 46: 47-66. [CrossRef]
44. Ashoor AA, Prochazka Jr T. Saudi Arabic speech audiometry. Audiology 1982; 21: 493- 508. [CrossRef]