Modified Rhyme Test for Evaluating Turkish Speech Intelligibility

Türkçe Konuşma Anlaşılabilirliğinin Değerlendirilmesi için Değiştirilmiş Kafiye Testi

Umut Arıöz[1], Banu Günel[2]

[1] Computer Engineering, Engineering Faculty, Gazi University, Ankara, Turkey
[2] Informatics Institute, Middle East Technical University, Ankara, Turkey

ABSTRACT

Objective: There are both objective and subjective tests for evaluating the speech intelligibility. Speech Intelligibility Index (SII) and Modified Rhyme Test (MRT) are the most common methods for objective and subjective tests, respectively. While objective tests do not depend on the language, subjective tests are applied to the hearing-impaired subjects in their language. In these tests, hearing-impaired subjects try to recognize the sound by looking at the word lists in their hand. The word lists can be presented as two-word or six-word choices in design. However, the preferred type a is six-word choice for MRT. Up to now, these word lists were developed in different languages with different methods. For Turkish language, only two-word choice lists were developed; there was no study related to the Turkish language with six-word choice. In this study, we aimed to develop applicable six-word Turkish MRT lists by using hearing-impaired subjects.
Methods: In this study, three different six-word Turkish MRT lists were developed according to the phonetic characteristics of the Turkish language, and these lists were tested via 12 hearing-impaired subjects. In the lists, 5 different Turkish phonetic characteristics were taken into account. For the analysis, the detailed audiograms (18 octave band) of hearing-impaired subjects were used.
Results: Detailed statistical analysis for phonetic variations were shown for each hearing-impaired subject. Having correct structure and robust word choices for MRT lists were shown by comparing the equivalent speech spectrum levels of sounds and MRT results of the hearing-impaired subjects.
Conclusion: New Turkish MRT lists were developed according to the phonetic characteristics of the Turkish language. Usability of the lists was shown by audiological test with the hearing-impaired subjects. Statistical analysis and the results of tests showed that developed Turkish MRT lists can be used in audiological studies as a reference material.

Key Words: Modified rhyme test; phonetic analysis; speech intelligibility; Turkish language

Received: 11.13.2014 **Accepted:** 01.23.2015

ÖZET

Amaç: Konuşma anlaşılabilirliğinin değerlendirilmesinde hem objektif hem de subjektif testler mevcuttur. Bu testler arasında en yaygın kullanılan metotlar objektif ve subjektif testler için sırasıyla Konuşma Anlaşılabilirlik İndeksi (KAT) ve Değiştirilmiş Kafiye Testi (DKT)'dir. Objektif testler uygulanan dilden bağımsız olmasına karşın subjektif testler uygulanan işitme kaybına sahip kişilerin kendi ana dillerinde uygulanmaktadır. Bu testlerde işitme kaybına sahip kişiler duydukları sesleri ellerindeki kelime listelerine bakarak tanımaya çalışmaktadırlar. Kelime listeleri iki kelime seçenekli veya altı kelime seçenekli olarak dizayn edilmektedirler. Fakat genel olarak DKT için altı kelime seçenekli kelime listeleri tercih edilmektedir. Şimdiye kadar bu kelime listeleri farklı metotlar kullanılarak farklı diller için geliştirilmiştir. Türkçe dili için sadece iki kelime seçenekli kelime listeleri geliştirilmiş; altı kelime seçenekli kelime listeleri için herhangi bir çalışma mevcut değildir. Bu çalışmada işitme kaybına sahip kişiler kullanılarak uygulanabilir altı kelime seçenekli Türkçe kelime listeleri geliştirmek amaçlanmıştır.
Yöntemler: Bu çalışmada Türkçe dilinin fonetik özellikleri dikakte alınarak üç farklı altı kelime seçenekli Türkçe DKT listesi geliştirilmiştir. Bu kelime listeleri 12 işitme kaybına sahip kişi kullanılarak test edilmiştir. Bu listelerde 5 farklı Türkçe dilinin fonetik özellikleri dikakte alınmıştır. Analizlerde işitme kaybına sahip kişiler için 18 oktav bandını içeren ayrıntılı odyogramlar kullanılmıştır.
Bulgular: Her bir işitme kaybına sahip kişi için fonetik değişimleri gösteren ayrıntılı istatistiksel analizler yapılmıştır. DKT listelerinin yapısal olarak doğru kurgulanması ve tutarlı kelime seçeneklerinin olması, seslerin eş konuşma spektrum seviyeleri ile işitme kaybına sahip kişilere uygulanan DKT sonuçlarının karşılaştırılmasıyla gösterilmiştir.
Sonuç: Yeni Türkçe DKT listeleri Türkçe dilinin fonetik özelliklerine göre geliştirilmiştir. Bu listelerin kullanılabilirliği işitme kaybına sahip kişilere uygulanan odyolojik testler ile gösterilmiştir. Istatistiksel analizler ve test sonuçları ile geliştirilen Türkçe DKT listelerinin odyolojik çalışmalarda referans materyali olarak kullanılabileceğini ortaya konmuştur.

Anahtar Sözcükler: Değiştirilmiş Kafiye Testi; fonetik analiz; konuşma anlaşılabilirlik; Türkçe dili

Geliş Tarihi: 13.11.2014 **Kabul Tarihi:** 23.01.2015

Address for Correspondence / Yazışma Adresi: Umut Arıöz, PhD, Gazi University Engineering Faculty Computer Engineering Room No: 116 Eti Mh. Yükseliş St. No: 5 Maltepe, Ankara, Turkey Tel: + 90 312 582 3116, E-mail: umut.arioz@gazi.edu.tr

doi:http://dx.doi.org/10.12996/gmj.2015.25

INTRODUCTION

Speech intelligibility measures are used mainly for evaluating synthetic or processed speech. Text to speech (TTS) synthesis systems and hearing loss simulation (HLS) are examples of the main fields for speech intelligibility measures (1-3). TTS systems try to convert normal language text into speech. These systems convert the text into symbols (the text normalization part) and, after applying linguistic analysis to them, convert these symbolic representations into sounds. HLS is also used as an important and useful research tool in order to understand the nature and underlying factors of hearing impairments. The primary goal of HLS studies is to model the impaired hearing of individuals, thereby eliminating the need to involve subjects with hearing loss in listening tests. For evaluating the effectiveness of the HLS, both subjective and objective speech intelligibility measures exist. The developed objective speech intelligibility indexes include the articulation index (AI) (4), speech transmission index (STI) (5) and speech intelligibility index (SII) (6). The most widely used subjective speech intelligibility measures are the diagnostic rhyme test (DRT) (7), modified rhyme test (MRT) (2) and mean opinion score (MOS) (8).

Studies on objective measurement methods for speech intelligibility commenced in Bell Laboratories in 1940. In 1969, the American National Standards Institute (ANSI) developed the articulation index (AI) (4). While calculating this index, the spectrum is divided into one-octave or one-third octave bands. The signal to noise ratio (SNR) values are then calculated for each band, and weighting factors are applied to the SNR values according to the importance of the particular frequency band. After normalization, the articulation index generates a value between zero and one; zero indicates completely unintelligible speech and one indicates completely intelligible speech.

Another objective measure, the speech transmission index (STI), was proposed by Steeneken and Houtgast (1980)(5). Calculation of STI is based on the preserved spectral differences of the phonemes. To obtain this index, an artificial input signal is constructed in place of a speech signal. For the construction and analysis of this artificial signal, a modulation transfer function is used. This transfer function is determined according to both the one octave bands of noise (125 Hz – 8 kHz) and the number of modulation frequencies (63 Hz-12.5 kHz). The analysis is based on specifying the significant SNR values of the artifical signal's octave bands. The general usage areas of STI are in evaluations of reverberation, non-linear distortions, noise, and echoes (IEC 60268-16, 1998).

The third objective measure, the speech intelligibility index (SII) was defined by the ANSI in 1997 (6). It has the same principle as the articulation index, but a number of corrections and a different weighting function for each frequency band were added to the SII calculations.

The definition of SII in the ANSI standard is the "product of band importance function and band audibility function, summed over the total number of frequency bands in the computational method" (6). The details of the computational method are explained in this standard.

For calculating the SII, there are four basic parameters: equivalent speech spectrum level, equivalent noise spectrum level, equivalent hearing threshold level, and band importance function. The SII generates non-linear results between 0 (unintelligible) and 1 (excellent intelligible), like the AI (9).

Subjective speech intelligibility tests are also used as commonly as objective speech intelligibility tests are. These tests can be administered using an open or a closed response methods. In the open response method, subjects try to recognize what they heard without having a word list to prompt them. Open response methods particularly use nonsense words. In using this method, better discrimination between words can be obtained for specific hearing conditions. However, the open response method requires long training sessions. In the closed response method, subjects attempt to select the words that they have heard from a prepared word list. Because of this property, the closed response method has the advantage of having a simple training session. The diagnostic rhyme test (DRT) and modified rhyme test (MRT) are examples of the closed response method.

The first study of the rhyme test was conducted by Fairbanks (1958) (10); inspired by this study, MRT was designed by House et al. in 1965 (2). In general, a list of twenty-five or fifty word groups with six rhyming words in each group are used in the MRT. In rhyming tests, the subjects try to choose the correct spoken word from among the group of written words that rhyme with it. The words in the groups are generally designed to have the same first character or same last character. There are MRTs constructed for different languages by considering the phonetic characteristics of that language (2,11,12). However, there has not been any study related to MRT for the Turkish language. Up to now, only DRT has been developed for the Turkish language (13).

The Diagnostic Rhyme Test (DRT) was developed by Voiers (1977) (7) and is based on only one pair of rhyming words. DRT has a simpler training session than the training for MRT subjects. Several DRT studies for evaluating the intelligibility of the Turkish language exist in the literature (14, 13).

Mean Opinion Score (MOS) is slightly different from the rhyme tests, and is used for speech coding algorithms and synthesized speech. Evaluation is performed according to the speech quality by using a questionnaire. Questions about overall impression, voice pleasantness, pronunciation, speaking rate, acceptance, and articulation are asked in MOS tests (8, 15).

The equivalent speech spectrum level (ESSL) is one of the different presentations of the speech spectrum. It has been in use since 1997 (6), and was developed by members of the ASA Working Group S3-79 (www.sii.to). According to the definition of ESSL of Working Group S3-79, "The Equivalent Speech Spectrum Level" is the speech spectrum level at the center of the listener's head that produces in an average human with uncovered ears an eardrum speech level equal to the eardrum speech level actually present in the listening situation to be analyzed. Before the SII can be applied to a given listening situation, the corresponding Equivalent Speech Spectrum Level must be derived. For example, when speech is presented over inserted earphones, only the speech spectrum level within the eardrum is known. Using the inverse of the freefield-to-eardrum transfer function (Table 3 of (6)) this eardrum level must be "projected" into the freefield, yielding the Equivalent Speech Spectrum Level. Different studies have been conducted by using the ESSL in speech processing applications (16, 17, 18).

The common problems for all subjective speech intelligibility tests are the long duration of the tests, the costs to arrange and carry out the tests, and the difficulty of finding suitable subjects.

In this study, three MRT lists, each containing six words, were prepared according to the phonetic characteristics of the Turkish language. All chosen words were recorded using native Turkish speakers. For assessment of the usability of the constructed word groups, an MRT was designed and applied to twelve hearing impaired subjects. Subjects listened to the prepared recordings in an isolated room. The results of the MRT were compared with the equivalent speech spectrum levels of recordings for each different phonetic characteristic.

MATERIAL AND METHOD

Subjects

Twelve hearing-impaired subjects, ten male and two female, participated in the study. Audiogram values (18 octave band) of the hearing impaired subjects can be seen in Table 1. These subjects were deliberately selected with a wide range of different types of hearing loss in order to obtain the unbiased effects of prepared methods and tests. Among these subjects, we have subjects who have mild hearing loss (subject 4), moderate hearing loss (subject 2, subject 7, subject 9, subject 10, subject 11), severe hearing loss (subject 1, subject 3, subject 5, subject 8), and profound hearing loss (subject 6, subject 12). For the study, approvals were gained from the Ethics Committee of Middle East Technical University. Each participant signed an informed consent form before participating in the study.

Table 1: Hearing threshold levels for 18 octave band for all subjects (Subjects vs. 18 octave band frequencies)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	160	**200**	**250**	**315**	**400**	**500**	**630**	**800**	**1000**	**1250**	**1600**	**2000**	**2500**	**3150**	**4000**	**5000**	**6300**	**8000**
Subject 1	45	45	45	46	48	50	51	53	55	56	58	60	61	63	66	75	84	85
Subject 2	10	10	10	10	10	10	10	10	11	16	25	34	41	49	59	65	71	78
Subject 3	30	30	30	30	30	30	29	27	26	32	43	54	61	69	79	83	88	93
Subject 4	15	15	15	14	12	10	10	10	10	10	10	11	15	21	30	37	46	53
Subject 5	45	45	46	51	60	68	71	73	75	77	81	85	86	88	90	92	96	99
Subject 6	70	70	70	72	76	80	83	86	90	92	96	99	100	100	100	100	100	100
Subject 7	40	40	40	42	46	50	55	62	69	71	73	75	74	72	70	68	66	69
Subject 8	25	25	26	29	34	39	43	46	49	50	50	51	57	67	78	82	88	97
Subject 9	30	30	31	34	39	44	49	54	58	56	51	46	45	45	44	40	39	51
Subject 10	55	55	55	55	55	55	54	52	51	51	53	55	54	52	52	60	67	62
Subject 11	33	36	39	40	40	40	40	40	40	41	43	45	45	45	45	47	50	50
Subject 12	20	20	20	20	20	20	23	26	31	43	63	83	91	99	108	110	110	110

Designing the MRT Lists

In the study, MRT lists and other methodological steps were designed with respect to previous similar studies (11, 12). These arrangements were, however, adopted according to the properties of the Turkish language. Three MRT lists were prepared with commonly used words to achieve better modeling of Turkish according to its phonetic characteristics: two 25-word group lists (in which the words started with the same character and ended with the same character) and one 50-word group list. The 50-word group list consisted of a number of sub-tests in which each sub-test measures the subject's ability to use acoustic information along different dimensions such as nasality, sustention, sibilation, compactness, and graveness. These dimensions were determined according to the phonetic characteristics of the Turkish language (14) (Figure 1). As can be seen from Figure 1, every Turkish consonant has distinct phonetic characteristics. The most relevant characteristics for the Turkish consonants are voicing (12 consonants) and sustention (10 consonants); the consonants which have the most phonetic characteristics are j and v (4 phonetic characteristics).

	Turkish Consonants																			
	b	**c**	**ç**	**d**	**f**	**g**	**h**	**j**	**k**	**l**	**m**	**n**	**p**	**r**	**s**	**ş**	**t**	**v**	**y**	**z**
	b	d͡ʒ	ç	ɗ	f	ɠ	ɦ	ʝ	k	ɭ	m	n	ɸ	ɾ	s	ʃ	θ	ⱱ	ʎ	ʐ
Nasality											+	+								
Sustention					+		+	+		+				+	+	+		+	+	+
Sibilation		+	+		+			+							+	+		+		+
Compactness		+				+		+	+							+			+	
Graveness	+				+						+		+					+		
Voicing	+	+		+		+		+		+	+	+		+				+	+	+

Figure 1: Phonetic characteristics of the Turkish consonants

To provide better understanding of these dimensions, articulation properties should be specified. Phoneticians classify the articulation of consonants in two groups. One group represents the location of articulation showing the placement of consonants, which starts from the lips and goes to the glottal region in the vocal tract. The other group is defined by the manner of articulation, which includes several factors about speech articulation such as the degree of narrowing of the vocal tract, raised or lowered position of velum, and being voiced or voiceless. Phonetic characteristics of Turkish and their articulation properties are shown in Figure 2 (19). This figure was prepared by combining all relevant literature information with our own efforts. Having more than one manner or place of articulation shows the wide range of Turkish consonants' positions. The first word shows the position that occurred most frequently in articulation. This figure was designed for comparing the properties of phonetic characteristics with the articulation of Turkish consonants, rather than comparing the Turkish consonants one by one. According to the figure, one can say that plosive Turkish consonants and Turkish consonants in the region of palatal and bilabial show more phonetic characteristics than others.

	Manner of Articulation			**Place of Articulation**			
Voicing	Fricative	Plosive	Voiced	Palatal	Bilabial	Postalveolar	
Nasality	Nasal	Voiced		Bilabial	Palatal		
Sustention	Plosive	Fricative	Voiced	Lablodental	Alveolar	Palatal	Pharyngeal
Sibilation	Plosive	Fricative	Voiced	Palatal	Dental	Alveolar	
Compactness	Plosive	Voiced	Voiceless	Velar	Palatal		
Graveness	Fricative	Voiced		Bilabial			

Figure 2: Properties of phonetic characteristics of the Turkish language according to the manner and place of articulation of sounds

Constructing the MRT Lists

Although there are six phonetic characteristics, five of them (all but voicing) were selected for constructing the 50-word group list to obtain meaningful statistics (10 words form 5 phonetics). Voicing was extracted because of the hardness of constructing a rhyme test with six words in Turkish, because the design of each of phonetic property was constructed with one phonetically-related word and five phonetically-unrelated words.

On the other hand, most of the consonants in Turkish have voicing property (Figure 1). Thus, five phonetically-unrelated words could not be found for the voicing property to accomplish the entire test design.

The suitable consonant vowel consonant (CVC) word groups were selected and three different MRT lists were constructed according to the rules of the design phase. In the first list, while all words have the same first character, the rest of the words begin with different characters (Table 2). In the second list, while all words have the same last character, the rest of the words end with different characters (Table 3). In the third list, while one of the words has a specific phonetic characteristic, the other words have different phonetic characteristics (Table 4).

Table 2: MRT list of 25 word groups with same first character

No	1. Word	2. Word	3. Word	4. Word	5. Word	6. Word
1	yat	yaz	yay	yan	yar	yas
2	fan	fas	far	fay	fal	faz
3	hap	hak	hat	haz	hal	han
4	kın	kıl	kıt	kın	kız	kış
5	bar	baz	bam	bal	bay	bas
6	güç	gül	gün	gür	güz	güm
7	kek	ker	kem	kel	kez	keş
8	sos	sol	son	sop	soy	sor
9	hal	hak	har	hap	hat	haz
10	baş	bas	bay	bal	baz	bam
11	yar	yan	yas	yat	yay	yaz
12	kez	kel	kem	kek	keş	ket
13	sor	sol	som	sos	soy	son
14	taç	tam	tav	tay	tan	tas
15	hak	ham	haz	hat	hal	hac
16	keş	kek	kem	kel	kez	ker
17	sop	soy	sol	som	son	sos
18	tay	tam	tak	tan	taç	tat
19	hal	hat	has	ham	har	hak
20	kıl	kız	kıt	kış	kın	kın
21	tak	tam	taç	tas	tan	tat
22	ham	hak	har	han	hap	hat
23	sap	saç	sac	sam	saz	san
24	şad	şal	şak	şam	şan	şap
25	kam	kaç	kaş	kan	kap	kat

Table 3: MRT list of 25 word groups with same last character

No	1. Word	2. Word	3. Word	4. Word	5. Word	6. Word
1	şen	sen	ben	ten	gen	yen
2	mal	dal	hal	fal	lal	şal
3	yay	vay	fay	hay	tay	çay
4	gaz	haz	baz	kaz	naz	yaz
5	tak	hak	kak	şak	pak	yak
6	sal	bal	dal	mal	şal	fal
7	dar	bar	kar	zar	far	nar
8	bor	mor	kor	zor	lor	hor
9	sur	dur	kur	nur	tur	vur
10	tat	hat	kat	yat	zat	mat
11	tam	cam	dam	gam	nam	ham
12	kil	çil	dil	fil	mil	pil
13	doz	boz	yoz	koz	poz	toz
14	pin	bin	din	kin	hin	tin
15	pim	çim	kim	mim	sim	tim
16	çan	şan	fan	han	san	zan
17	dar	bar	gar	far	var	zar
18	zan	çan	fan	han	tan	yan
19	fit	bit	çit	kit	hit	sit
20	sol	bol	hol	rol	yol	dol
21	kel	bel	jel	gel	tel	yel
22	hat	zat	kat	mat	tat	yat
23	ten	ben	gen	men	fen	yen
24	rol	bol	gol	hol	mol	sol
25	bor	zor	hor	kor	lor	mor

Table 4: MRT list of 50 words groups according to their Turkish phonetic features (Words with phonetic property were placed in the first order)

No	1. Word	2. Word	3. Word	4. Word	5. Word	6. Word	Feature
1	**ben**	sen	şen	ten	gen	yen	graveness
2	**fal**	dal	hal	mal	lal	şal	graveness
3	**mol**	gol	sol	hol	rol	kol	graveness
4	**put**	tut	dut	kut	şut	gut	graveness
5	**vay**	çay	fay	hay	tay	yay	graveness
6	**baz**	haz	gaz	kaz	naz	yaz	graveness
7	**pak**	hak	kak	şak	tak	yak	graveness
8	**ver**	ger	ker	şer	ter	yer	graveness
9	**kem**	ker	keş	kel	kek	kez	graveness
10	**sop**	son	sol	sor	sos	soy	graveness
11	**mal**	bal	dal	sal	şal	fal	nasality
12	**mil**	kil	pil	zil	dil	fil	nasality
13	**nar**	bar	kar	zar	far	dar	nasality
14	**mor**	kor	bor	zor	lor	hor	nasality
15	**naz**	kaz	faz	yaz	baz	saz	nasality
16	**nem**	dem	gem	yem	kem	hem	nasality
17	**nur**	dur	kur	sur	tur	vur	nasality
18	**mat**	hat	kat	yat	zat	tat	nasality
19	**kem**	kez	kel	kek	ket	keş	nasality
20	**son**	som	sol	sor	sos	soy	nasality

22	**fil**	çil	dil	kil	mil	pil	sustention
23	**fas**	pas	kas	yas	tas	bas	sustention
24	**far**	bar	çar	kar	nar	dar	sustention
25	**yoz**	boz	doz	koz	poz	toz	sustention
26	**sek**	bek	dek	kek	pek	tek	sustention
27	**hin**	bin	din	kin	pin	tin	sustention
28	**jul**	bul	çul	dul	kul	pul	sustention
29	**sim**	çim	kim	mim	pim	tim	sustention
30	**şal**	şak	şam	şad	şan	şap	sustention
31	**şan**	çan	fan	han	san	zan	compactness
32	**gar**	bar	dar	far	var	zar	compactness
33	**şark**	fark	bark	park	sark	çark	compactness
34	**gen**	ben	fen	men	sen	ten	compactness
35	**yan**	çan	fan	han	tan	zan	compactness
36	**kit**	bit	çit	fit	hit	sit	compactness
37	**yol**	bol	hol	rol	sol	dol	compactness
38	**kim**	çim	mim	pim	sim	tim	compactness
39	**bay**	bam	bas	baz	bal	bar	compactness
40	**hak**	hat	han	haz	hap	hal	compactness
41	**kez**	kem	kek	kel	ker	keş	sibilation
42	**kız**	kıt	kıl	kın	kış	kır	sibilation
43	**jel**	bel	kel	gel	tel	yel	sibilation
44	**zam**	bam	dam	ham	nam	tam	sibilation
45	**zat**	hat	kat	mat	tat	yat	sibilation
46	**fen**	ben	gen	men	ten	yen	sibilation
47	**var**	bar	dar	yar	gar	kar	sibilation
48	**sol**	bol	gol	hol	mol	rol	sibilation
49	**çin**	bin	din	hin	kin	pin	sibilation
50	**zor**	bor	hor	kor	lor	mor	sibilation

Stimuli

Chosen words in the MRT lists were spoken by a male and a female native Turkish speaker and recorded in an acoustically treated studio. The speakers repeated the sentence "Aşağıdakilerden <Word> kelimesini seçer misiniz?" which can be translated into English as "Could you choose the word <Word>?" This carrier phrase was chosen to get more concentration on the rhyme test words of the subjects. This phrase was explained to the subjects, who were requested to pay attention to the rhyme test words before the test. This is the common method of rhyme tests for similar studies. The speech recordings were obtained at the 48000 Hz sampling frequency with a 16 bit resolution, using a Sennheiser M64 pre-polarized condenser microphone and EDIROL UA-1000 Audio Capture device.

MRT Test Design

A printed copy of the MRT lists was given to each of the subjects. Subjects were then asked to mark the word they heard for each spoken sentence. All subjects listened to the same MRT list sets; however, the orders of the words were randomized to eliminate the effect of word order on the results. Hearing-impaired subjects listened to the stimuli using headphones in a noise-free environment. Subjects were asked to remove their hearing aids during the tests. All subjects were trained until they got used to the MRT procedures. Speech recordings were only provided to the selected ear of each hearing impaired subject. Tests generally lasted 15-20 minutes per subject.

RESULTS

In the phonetic analysis, the equivalent speech spectrum levels of recordings, box-plot results, and percentages of correct choices of MRT for each phonetic group were calculated for all subjects. For objective comparison of performance, equivalent speech spectrum levels of the recordings were calculated and used together with the hearing thresholds of subjects according to their 18 octave band-audiograms.

Equivalent Speech Spectrum Levels (ESSLs)

ESSL is calculated by taking the logarithms of root mean squares of each frequency band. The results are time-averaged values for input speech. To obtain more compact result, the results were shown according to the average values for males and females, as within-gender properties show similar characteristics for ESSL. The unit is decibel (dB) (6). The ESSLs for each phonetic characteristic for both genders were calculated according to the standard 18 octave bands in the unit of dB (Figure 3).

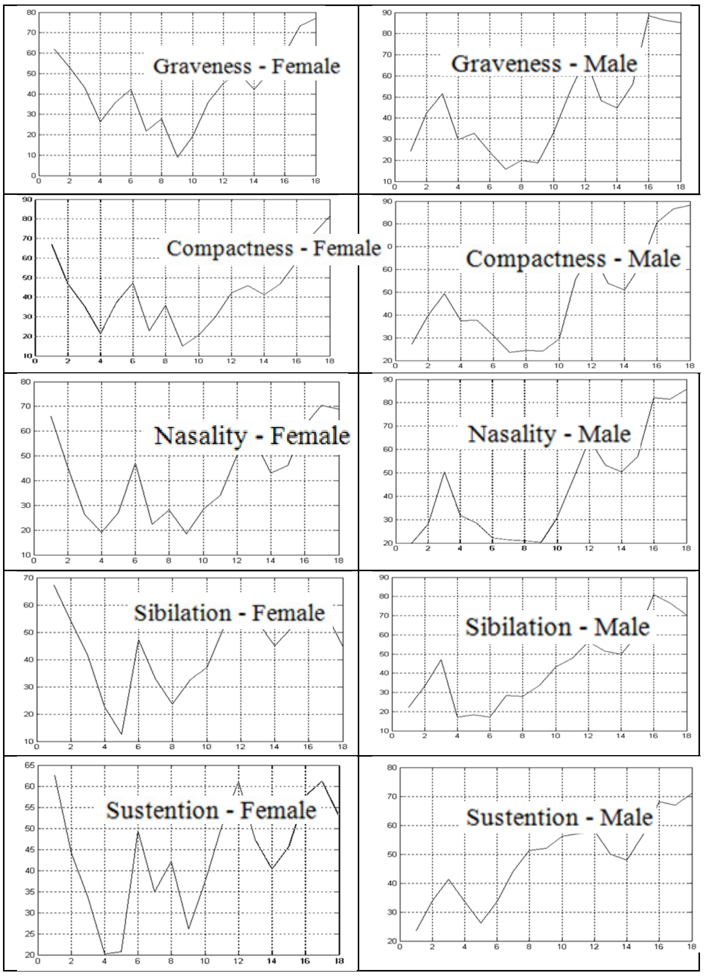

Figure 3: Equivalent speech spectrum levels of each phonetics (18 octave bands versus dB)

The Box-plots

The percentages of correct choices of MRT according to the each phonetic are presented for each subject as a box-plot in the Figure 4 (for the results from Subject 1 to Subject 6) and Figure 5 (for the results from Subject 7 to Subject 12). By using these box-plots, minimum value, maximum value, median value, and range of variation (from first quartile to third quartile) of each subject's MRT result can be easily seen.

For Subject 1, the median values for all phonetics are very close to each other (~80%) and each phonetic has small standard deviation. Hence, the smallest standard deviation was obtained for the sustention property (12%). For Subject 2, all phonetics have a high range of variation and the median value for each phonetic is different from the others (mean is approximately 65%). For Subject 3, most of the median values are equal to the minimum or maximum of the MRT percentages. These results make up the smaller range of variation for related phonetics (sustention, sibilation, nasality and graveness).

For Subject 4, all phonetics except graveness have a small range of variation, but have different mean and median values. For Subject 5, the results of MRT percentages are lower than 65%, especially sustention and sibilation properties are in the range of 15% and 40%. For Subject 6, the results of MRT percentages show the variability among the phonetics. While the median value of the sustention property is 30%; the median value of the nasality property is 76%.

For Subject 7, the median values for all phonetics are very close to each other (~60%) and the standard deviations of all phonetics are similar (between ~40% and ~80%). For Subject 8, each phonetic shows different characteristics across all values. However, all phonetics except the sibilation property have small standard deviations. For Subject 9, all phonetics have the minimum of 60% MRT percentages with higher median values. The average median values for all phonetics is approximately 80%. For Subject 10, most of the median values are equal to the minimum or maximum of the MRT percentages. These results construct the smaller range of variation for related phonetics (sustention, sibilation, nasality and graveness), as seen in Subject 3. For Subject 11, the average median value for all phonetics is approximately 75% and all phonetics have the minimum percentage 65% in the range of variation. For Subject 12, MRT percentages are greater than 60% for all phonetics except sustention. The sustention properties have a 35% median value and all results are below the 50% MRT percentage.

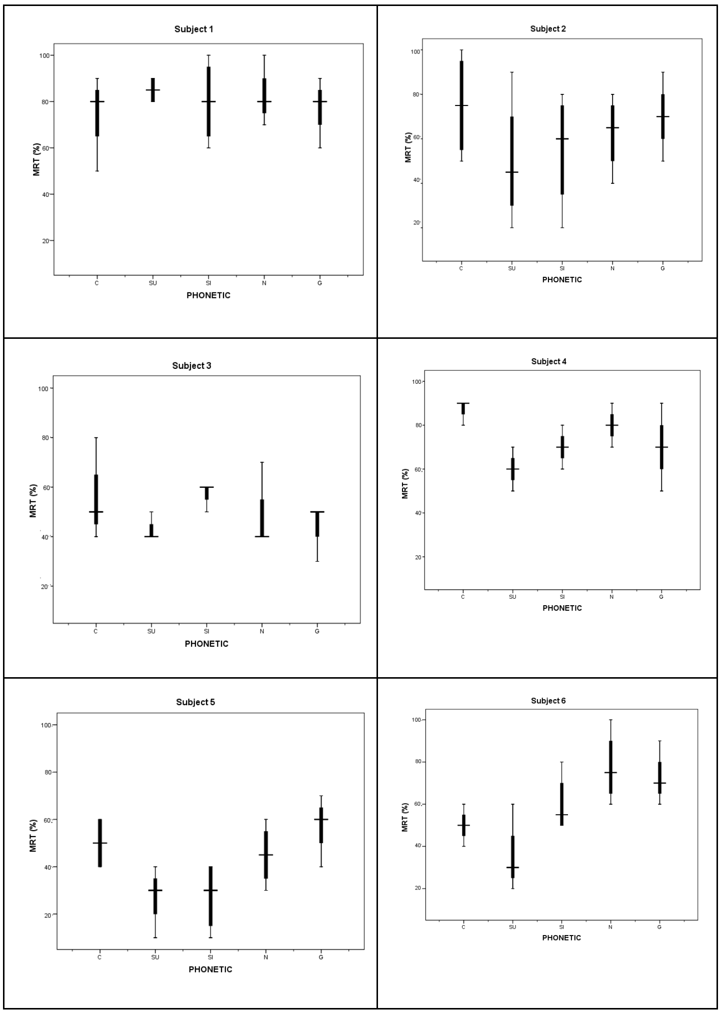

Figure 4: The box-plot results of MRT percentages of each phonetic for the Subject 1 to Subject 6 (C: Compactness, SU: Sustention, SI: Sibilation, N: Nasality, G: Graveness)

DISCUSSION

The meaningful comparison between ESSLs and hearing threshold levels is made especially for the high frequency region, because dB levels of phonetics are dominant in that region. According to the 18 octave band hearing threshold levels, five subjects (Subject 3, Subject 5, Subject 6, Subject 8, and Subject 12) have hearing loss greater than 90 dB in the high frequency region. This means that those subjects had greater difficulty in understanding the words within this region. The box-plot results of MRT percentages of each phonetic support this statement. In particular, Subject 3, Subject 5 and Subject 6 have a mean percentage lower than 50%. Although Subject 8 and Subject 12 have hearing loss greater than 90 dB in the high frequency region, they have a mean percentage higher than 50%. This situation can be explained by a subject having very low hearing thresholds in the lower frequency regions. These low thresholds can affect the understanding of a greater number of the words tested, because speech is in the range of 1 kHz and 4 kHz. On the other hand, although Subject 10 has a hearing loss lower than 90 dB in the high frequency region, he/she has a mean percentage lower than 50%. In contrast with the previous subject, this condition can be explained by having a very high hearing threshold in the lower frequency regions, which can affect the understanding of a lower number of the words tested for Subject 10.

According to the ESSLs, having higher dB amounts can easily be seen in the high frequency region (octaves between 12 and 18) for both female and male recordings. This shows that words in the MRT lists have mostly high frequency contents. Among the phonetics, while sibilation and sustention have the minimum dB contents, graveness and compactness have the maximum dB contents for both female and male recordings. These ESSLs are consistent with the results of MRT percentages of each phonetic, as sustention is the lowest MRT percentage for 8 subjects (Subject 2, Subject 3, Subject 4, Subject 5, Subject 6, Subject 9, Subject 10, and Subject 12) and sibilation is the lowest MRT percentage for 3 subjects (Subject 1, Subject 5, and Subject 8).

According to these results, the hearing thresholds of the subjects, ESSLs of the recordings, and MRT percentages registered results consistent with respect to each other. This consistency demonstrates the correct structure and word choice of the MRT lists developed for this analysis.

CONCLUSION

Three different MRT lists were developed for the Turkish language. These lists were tested on hearing-impaired subjects with MRT. Using phonetic analysis, the effects of different phonetics and usability of the lists were demonstrated with respect to MRT results.

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

REFERENCES

1. Pisoni DB, Hunnicutt S. Perceptual evaluation of MITalk: The MIT unrestricted text-to-speech system, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing; New York: IEEE Press, 1980; 5: 572 -5.

2. House AS, Williams CE, Hecker MHL, Kryter KD. Articulation testing methods: consonantal differentiation with a closed response set. Journal of the Acoustical Society of America. 1965;37: 158–66.

3. Villchur E. Simulation of the effect of recruitment on loudness relationships in speech, Journal of the Acoustical Society of America. 1974;56:1601–11.

4. ANSI-S3.5 (1969). American National Standard methods for the calculation of the Articulation Index, American National Standards Institute, Inc.: New York.

5. Steeneken H, Houtgast T. A physical method for measuring speech transmission quality. Journal of the Acoustical Society of America, 1980;67: 318-26.

6. ANSI-S3.5 (1997). American National Standard methods for the calculation of the Speech Intelligibility Index. American National Standards Institute, Inc.: New York.

7. Voiers WD. Speech Intelligibility and Speaker Recognition. In: Hauley, Dowden (Eds.), Hutchinson and Ross Inc. 1977.

8. ITU-T Recommendation P.85. Telephone transmission quality subjective opinion tests. A method for subjective performance assessment of the quality of speech voice output devices.1994.

9. Sherbecoe RL, Studebaker GA. Audibility-index functions for the connected Speech Test. Ear & Hearing, 2001;23: 385-98.

10. Fairbanks G. "Test of phonemic differentiation: The rhyme test", J.Acoust. Soc.Am. 1985;30:596-600.

11. Li Z, Tan EC, McLoughlin I. Teo TT. Proposal of standards for intelligibility tests of Chinese speech, IEE Proc. on Vision, Image and Signal Processing, 2000; 147:254-60.

12. Tihelka D. Matoušek J. The Design of Czech Language Formal Listening Tests for the Evaluation of TTS Systems, In Proceedings of Fourth International Conference on Language Resources and Evaluation - LREC 2004, Lisbon, Portugal, vol VI, 1099-2002.

13. Sak H, Güngör T, Safkan Y. A corpusbased concatenative speech synthesis system for Turkish. Turkish Journal of Electrical Engineering, and Computer Sciences, 2006;14: 209-23.

14. Palaz H, Bicil Y, Kanak A, Dogan MU. New Turkish Intelligibility Test for Assessing Speech Communication Systems. Journal of Speech Communication, 2005; 47: 411-23.

15. Salza PL, Foti E, Nebbia L, Oreglia M. MOS and pair comparison combined methods for quality evaluation of text to speech systems. Acta Acustica, 1996;82: 650-6.

16. Royster HL. The Noise-Vibration Problem-Solution Workbook. ISBN: 1931504148. Aiha Press.2001.

17. Sauert B, Vary P. Near End Listening Enhancement Optimized With Respect To Speech Intelligibility Index. 17th European Signal Processing Conference (EUSIPCO 2009). Glasgow, Scotland, August 24-28.

18. Loizou PC. Speech Enhancement: Theory and Practice, Second Edition. Taylor & Francis Group, LLC. Press. 2013.

19. International Phonetic Association (IPA), (1999). Handbook of the International Phonetic Association—A Guide to the Use of the International Phonetic Alphabet, Cambridge University Press, Cambridge, UK.