

# The homogeneity of audibility and prosody of Zulu words for Speech Reception Threshold (SRT) testing

\*Seema Panday, #Harsha Kathard, #Mershen Pillay,

\*Cyril Govender

\*University of KwaZulu-Natal #University of Cape Town

## Abstract

The aim of this investigation was to determine which of 58 preselected Zulu words developed by Panday et al. (2007) could be used for Speech Reception Threshold (SRT) testing. To realize this aim the homogeneity of audibility of 58 bisyllabic Zulu low tone verbs was measured, followed by an analysis of the prosodic features of the selected words. The words were digitally recorded by a Zulu first language male speaker and presented at 6 intensity levels to 30 Zulu first language speakers (18 -25 years, mean age of 21.5 years), whose hearing was normal. Homogeneity of audibility was determined by employing logistic regression analysis. Twenty eight words met the criterion of homogeneity of audibility. This was evidenced by a mean slope of 50% at 5.98%/dB. The prosodic features of the twenty eight words were further analyzed using a computerized speech laboratory system. The findings confirmed that the pitch contours of the words followed the prosodic pattern apparent within Zulu linguistic structure. Eighty nine percent of the Zulu verbs were found to have a difference in the pitch pattern between the two syllables i.e. the first syllable was low in pitch, while the second syllable was high in pitch. It emerged that the twenty eight words could be used for establishing SRT within a normal hearing Zulu speaking population. Further research within clinical populations is recommended.

**Key Words** Speech Reception Threshold, Zulu, Zulu First Language Speakers, homogeneity of audibility, acoustic analysis, prosodic features

The selection of words for SRT testing in languages other than English has recently received much attention in the literature. This follows Nissen, Harris, Jennings, Eggett & Buck's (2005) assertion that it is essential that SRT testing be conducted in the language of the listener. Given this imperative and to attain valid and reliable SRT results, research has been conducted in many languages such as Arabic (Ashoor & Proshazka, 1985); Tiwi (Plant, 1990); Walpiri (Plant, 1990); Brazilian Portuguese (Harris, Goffi, Pedalini, Gygi & Merrill, 2001); Korean (Harris et al., 2003); Polish (Harris et al., 2004); Mandarin (Nissen et al., 2005) and Taiwanese Mandarin (Nissen, Harris & Dukes, 2008). However, SRT word lists in African languages such as Zulu have received very little or no attention although Zulu is spoken by almost 8.5 million people in South Africa (Grimes, 1992). Therefore, it is necessary that a linguistically-matched SRT word list be formulated according to specific criteria to enable the appropriate assessment of hearing of the many Zulu speaking South Africans, the majority of whom reside in the province of KwaZulu-Natal (KZN).

Panday, Kathard, Pillay & Govender (2007) have described the developmental process of adapting criteria for SRT word lists by considering Zulu linguistic structure. The adaptation of the original SRT criteria to suit the linguistic structure of the language, for

which the list is being developed, has received strong support in international literature. Evidence of the development of SRT word lists in other mother tongue languages is seen in Arabic word lists (Ashoor & Proshazka, 1985); Walpiri & Tiwi (Plant, 1990); Mandarin (Nissen et al. 2005) & Taiwanese Mandarin (Nissen et al. 2008). Panday et al. (2007) have reported on the identification and selection of fifty eight Zulu words that met the following linguistic criteria i.e. the words were bisyllabic, phonetically dissimilar, familiar, low tone verb imperatives. These criteria were adapted from the original SRT criteria as specified by Hudgins, Hawkins, Karlin & Stevens (1947) i.e. the words had to be familiar, bisyllabic, equally stressed words (spondees) that were phonetically dissimilar, and which had to be homogeneous for audibility. The focus of this paper is to describe which of the fifty eight preselected Zulu words met the criterion of homogeneity of audibility. In addition, a description of the prosodic features of the most homogenous words selected was made, so that the words finally selected from this process could be used for SRT testing within a selected Zulu first language speaking population.

Contact: Seema Panday

Discipline of Audiology University of KwaZulu –Natal Westville Campus

Durban South Africa

Email: [pandayse@ukzn.ac.za](mailto:pandayse@ukzn.ac.za)

Tel: 031-2607623

Homogeneity of audibility has been described as the ease at which the words are understood when spoken at a constant level of intensity (Silman & Silverman, 1991). It is regarded as one of the most important criteria for the selection of SRT test materials (Nissen et al., 2005). Homogeneity of audibility of words in a list can be achieved in two ways, i.e. by selecting only those words that reach the listener's ear at the same intensity or by recording the individual words in such a way that they all tend to be heard at the same level of production (Beattie, Svihohee & Edgerton, 1975). Schill (1985) explained that in order to meet the criterion of homogeneity of audibility, the percentage of words correctly recognized must increase rapidly with a relatively small increase in intensity. This is usually illustrated through the use of the performance-intensity curve (Brandy, 2002). The performance intensity curve is the more recent term used for the psychometric function which describes the relationship between some measure of performance and a stimulus (Kruger & Kruger, 1997).

The performance intensity curve serves to illustrate how well the speech sample (words in the list) is correctly identified as a function of various intensity levels. The steepness of the curve determines the precision with which the threshold can be obtained (Hudgins, Hawkins, Karlin & Stevens, 1947). Nissen et al. (2005) added that the steepness or slope of the performance intensity curve is an important factor to consider in selecting words for SRT testing. Hence, these aspects were considered when determining which of 58 Zulu words preselected by Panday et al. (2007) met the criterion of homogeneity of audibility.

However, obtaining ideal steep performance intensity curves together with homogeneity of audibility has posed serious challenges over the years. The literature seems to indicate that the number of words selected is reduced when the criterion for homogeneity of audibility is applied. This was seen when the original SRT word list in English was reduced from 84 words to 36 words (Hirsh, Silverman, Reynolds, Eldert & Benson, 1952) and was further reduced to 22 words (Bowling & Elpern, 1961).

Bowling and Elpern's (1961) findings were supported by research conducted by Curry and Cox (1966) who found similar results. The range of intelligibility between the words in the Curry and Cox (1966) study for normal hearers was 8%/dB compared to 10%/dB in the Bowling and Elpern (1961) study. The homogeneity of audibility of the words suggested in the above studies was also challenged by Beattie et al. (1975). These authors questioned whether the word differences noted in the Bowling and Elpern (1961) study were due either to inconsistencies in the recording process or due to the selection of the words.

Later, Young et al. (1982) explained these discrepancies by presenting a more stringent methodological framework for determining how words for SRT testing in English met the criterion of homogeneity of audibility. Young et al. (1982) postulated that

previous studies were confounded by the learning effect that occurred within the word lists. Bowling and Elpern (1961) and Beattie et al. (1975) tested the same randomized words at several intensity levels using a 2dB step increase in the intensity. Young et al. (1982) regarded this as a confounding factor as the participants were possibly performing better because of having learnt the words in the list rather than using hearing acuity to determine loudness or clarity of the words.

More recently, however, Nissen et al. (2005) utilized a detailed mathematical model to determine which of 138 Mandarin words met the criterion of homogeneity of audibility. These authors provided a motivation for the use of the logistic regression model for the measurement of homogeneity of audibility. The logistic regression model according to Nissen et al. (2005) provided a description of the slope at 50% of the performance intensity curves of each word, the estimated threshold of intensity at which each word was heard and individual performance intensity curves for the words. This allows the researcher to rapidly identify those words that are being heard easily compared to the words that were not (Nissen et al., 2005). The advantage of using this model is that the rate of change of intelligibility is accurate and available instantly.

In addition, Nissen et al. (2005) described favourable results with regard to the slope at 50% for Mandarin words. There was an almost equivalent performance with that of the English words. The 11.3 percent per decibel (%/dB) obtained in the Nissen et al. (2005) study was perhaps due to a digital adjustment made to the intensity of each of the words, so that the 50% of each word selected was equal to the mean pure tone average of the participants in the study. This adjustment controlled for variability of the performance intensity curves and thereby improved the overall homogeneity of audibility of the words.

Clearly, the measurement of homogeneity of audibility of the words for SRT testing is not a simple issue. Homogeneity of audibility of the words and the performance intensity function of each word are not only influenced by methodological and procedural variables as described earlier, but can also be influenced by several other factors that have been recently discussed in the literature. Some of these factors include, the type of stimuli used (Kruger & Kruger, 1997), the number of syllables in the words (Chetty, 1990), the recording (Nissen et al., 2005) and calibration of equipment (Lyregaard, 1997), the choice of the speaker used for the recording of the words and most importantly the structure of language.

It is important that test material consists of items familiar to the listener. The items must represent the real world of the listener (Chermack & Musiek, 1997). This contention appears relevant to the measurement of homogeneity of the words in Zulu. Therefore, apart from the acoustic-phonetic characteristics of the words, the

lexical processes, such as the semantic, syntactic and content of the words can also influence the way words are heard. Consequently, a "steeper performance intensity curve results when the task is simple, the stimuli are familiar, and most homogeneous" (Kruger & Kruger, 1997, p. 234). Therefore, within the present study the 58 preselected Zulu words were items that were most familiar in the Zulu language. This is also in keeping with the suggestions made by Borg, Wilson & Samuelsson (1998) who highlighted the importance of selecting familiar items for SRT testing. Panday et al. (2007) provide a detailed explanation of how the 58 preselected Zulu words met the criterion of familiarity.

Another factor that influences the homogeneity of audibility of the words is the number of syllables in the word. Bisyllabic words were traditionally used for SRT testing. However, the study conducted by Nissen et al. (2005) obtained steep performance intensity curves with trisyllabic words. One could argue that the increase in the number of acoustic cues in the trisyllabic words can affect the steepness of the curve because the increase in the acoustic cues available to the listener allows the word to be easily identified at lower intensities. Therefore, measurement of homogeneity of audibility in the present study could be influenced by the number of syllables in the Zulu words. However, given that there is no reference point for what constitutes an acceptable number of acoustic cues, it would be inappropriate to perform cross-linguistic comparisons between bisyllabic words versus trisyllabic words without a fuller semantic-syntactic linguistic analysis. A more critical factor influencing the validity of measurement of homogeneity of audibility, however, is whether the words are presented in recorded versus monitored live voice testing (Nissen et al., 2005).

Recorded and monitored live voice testing methods are both recommended for clinical use (American Speech-Language Hearing Association [ASHA], 1988; Brandy, 2002; Roeser, Valente & Horsford-Dunn 2000; Silman & Silverman, 1991; Stach, 1998). However, in terms of measuring homogeneity of audibility of the words, it is advised that recorded materials be used (Nissen et al., 2005). The use of recorded materials standardises the composition and presentation of the test words. Furthermore, the recorded method allows for uniform presentation of the test words (ASHA, 1988). Therefore, recorded material was chosen in the present investigation to ensure favourable performance intensity curves.

Moreover, researchers should be cognizant of the choice of speaker used in the recording of the stimuli. The stimuli must not only be presented in the language of the listener, but the speaker presenting the stimuli must be of a similar linguistic and dialectical background as the listener (Nissen et al., 2005). In light of this suggestion, the speaker in the present investigation was a male Zulu First Language speaker from KZN. A male speaker was

considered, as recent literature has indicated no clear gender differences influencing the homogeneity of audibility (Cambron, Wilson & Shanks, 1991; Nissen et al., 2005).

Thus, the results of homogeneity of audibility may or may not be influenced by the factors discussed above. However, the unique structure of the Zulu language may impact on the outcome of homogeneity of audibility of 58 preselected Zulu words and deserves some discussion and investigation.

The consonant system in Zulu is considered more complex than that of English. Apart from the known stops, fricatives, and approximants, Zulu also has three prominent click sounds (Doke, 1930). The click sounds are orthographically transcribed as /c/ (dental click); /q/ (palato-alveolar click); and the /x/ (lateral click) (Doke, 1930). These click sounds appear frequently in the vocabulary of Zulu. This has certain implications for the measurements of homogeneity of audibility of the words, i.e. inclusion of click phonemes in the word list may influence the variability in the performance intensity curves of the words that have click sounds. Click sounds have a higher spectral energy and this may result in improved audibility of the words that contain clicks. Therefore, the structural variation in the language and the fact that Zulu is also a tonal language (Rycroft & Ngcobo, 1979) could influence the measurement of homogeneity of audibility and pose a possible challenge when determining those words that are most homogenous.

In addition, the role of tone has three distinct functions in Zulu, i.e. semantic (affecting the meaning of words), grammatical, and emotional (Cope, 1982). These tonal variations may influence the meaning of the stimuli, thus having further implications for speech understanding and identification.

Tonal variations are lexically significant in Zulu (Rycroft & Ngcobo, 1979), implying that changes in the pitch of the word may correspond with changes in the meaning of the word. Hence, the introduction of acoustic analysis via pitch extraction and energy values is suggested in addition to evaluating the word for homogeneity of audibility. The inclusion of acoustic analysis would confirm the tonal patterns of the verbs in Zulu. Inherent in an acoustic analysis of words is the description of the prosodic features of a language. In considering the prosodic features of Zulu, we find that they are different from those of English. Therefore, an evaluation of the prosodic features of a language via appropriate acoustic analysis serves to improve the validity of the Zulu words selected for SRT testing.

Prosodic features of speech usually include length, accent, stress, tone and intonation (Fox, 2000). Furthermore, intonation refers to the patterns relating to the rise and fall, and the stress in a language (Kent & Read, 1992). In considering Zulu, we find that it is a "non-stress" language (Rycroft & Ngcobo, 1979) e.g. bisyllabic words such as "hamba" meaning 'go' in English, do not

have equally stressed syllables. Stress is not used to indicate emphasis nor is it used to differentiate words or syllables as seen in English (Cope, 1982). Instead, the prominence on syllables is often used to describe emphasis. This can be seen e.g. when low tone bisyllabic verbs are used in SRT testing, the lengthening of the vowel on the first syllable results in the second syllable gaining prominence.

Similarly, Plant (1990) reported that the prominence or stress of the trisyllables in Tiwi was on the penultimate syllable of the Tiwi words. Plant (1990) therefore, suggested an acoustic analysis of the words to confirm the prosodic patterns of the Tiwi words before the words were considered for inclusion in SRT testing.

Thus, it is strongly suggested that the prosodic features of the selected Zulu words also be analyzed acoustically. An acoustic analysis via pitch extraction and energy values of the vowels would serve to confirm the tonal patterns of the verbs selected i.e. to examine whether they are low tone vs high tone verbs. The words selected for SRT testing, therefore, would be homogenous in terms of the prosodic features of Zulu as well as for audibility.

The importance of acoustic analysis in SRT word list development, however, is lacking in the literature, but the above motivation together with the understanding of basic acoustics serves as a basis for the inclusion of acoustic analysis in the present investigation. Furthermore, Brandy (2002) indicates that the acoustic analysis of speech is important, as speech sound perception is dependant on both acoustic -phonetic factors and the higher order linguistic processing.

Contributions to the field with regard to selecting words for SRT testing in languages other than English are acknowledged. However, factors influencing the outcome of homogeneity of audibility together with the uniqueness of the Zulu language, suggest a need to determine which of the Zulu words proposed by Panday et al. (2007) meet the criterion of homogeneity of audibility and to describe the prosodic features of the most homogenous words.

Thus, the research question for this investigation was: Which of the 58 preselected Zulu words by Panday et al. (2007) could be used for SRT testing within a Zulu speaking population in KZN?

## METHOD

### AIM

The aim of the investigation was to determine which of the 58 preselected Zulu words could be used for SRT testing within a Zulu speaking population in KZN.

The following objectives were generated to achieve the above aim.

### OBJECTIVES

1. To determine which of the 58 preselected Zulu words met the criterion of homogeneity of audibility.
2. To describe the prosodic features, specifically pitch and en-

ergy contours, of the most homogenous words selected in objective 1

### RESEARCH DESIGN

A descriptive cross sectional design was adopted in this study (Connolly, C, Statistician, Medical Research Council, personal communication, August 2009). Such a design was considered suitable as it provided the data at a fixed point in time (Maxwell & Satake, 2006). A quantitative approach was followed because the study involved the analysis of responses to the audibility of the words at different intensity levels. This was followed by an analysis of the prosodic patterns of the Zulu words.

The methodological framework for each objective is described below.

### OBJECTIVE 1: TO DETERMINE WHICH OF THE 58 PRESELECTED ZULU WORDS MET THE CRITERION OF HOMOGENEITY OF AUDIBILITY

#### *Participant selection criteria*

The following participant selection criteria were adopted for this part of the study.

The participants had to be between the age of 18 - 25 years, as hearing sensitivity is judged to be at its peak during this age period (Jerger, 1970).

The participants had to speak Zulu as their first language and had to have been permanent residents of KZN for more than 10 years. This criterion ensured that the participants would be familiar with the items selected for testing and that there would be no confounding dialectal factors.

The participants could have been male or female, as gender specificity was not investigated.

The participants were required to have had no previous exposure to industrial or recreational noise, no exposure to ototoxic drug consumption and must have reported no previous medical, neurological or acquired illnesses. Further, participants had to have normal hearing. No family history of hearing loss should be reported. The above factors could contribute to auditory disorders that could confound the results. Bess and Humes (2003) describe the above factors as exogenous and endogenous factors, which could result from varying degrees of hearing loss. The above factors were evaluated using a case history questionnaire and were further confirmed by otoscopic examination, pure tone audiometry and immittance testing.

The participants had to have no exposure to hearing testing before their participation in this study (Robinson & Koenings, 1979). This criterion was necessary in order to control for any learned effects regarding any of the behavioral measures used.

#### *Sampling method*

Purposive non-probability sampling was used for this part of

the study. Zulu first language speakers attending the University of Kwa-Zulu Natal were targeted as the study population. Advertisements for participants included posters and an email advert on the University LAN. The advertisement included the specific criteria for participation in the study. Thirty-two participants responded to the original advertisements. However, once the sample selection criteria and the hearing testing process via case history, otoscopic examination, pure tone audiometry and immittance testing were completed, only thirty participants were considered suitable for the study. The two participants who did not meet the selection criteria were referred for further audiological testing and ENT management.

### **Participants**

Thirty normal hearing Zulu first language speaking (ZFLS) students from KZN were included as participants to realise this first objective of the study. The mean age of the participants was 21.5 years. There were 22 male and 8 female participants. Thirty adults were considered adequate, as the minimum number used for the evaluation of homogeneity of audibility in other studies was reported to be between twenty and twenty five (Nissen et al., 2005).

### **Materials and Apparatus**

#### **Words**

Fifty eight bisyllabic, low tone Zulu verbs were measured for homogeneity of audibility. The process resulting in the selection of the fifty eight words is described by Panday et al. (2007). Bisyllabic low tone verbs were selected according to careful consideration of Zulu linguistic structure. Furthermore, the 58 preselected words were subjected to analysis at two levels i.e. analysis in terms of commonality of the words in the Zulu language, a linguistic rating strategy in terms of familiarity and phonetic dissimilarity and tone (Refer to Panday et al., 2007).

#### **Recordings**

In order to achieve objective 1, 58 preselected words had to be recorded onto compact discs (CD). The recording was done in a recording studio. The equipment and the procedure for the recording was according to the guidelines stipulated by Lyregaard (1997); ASHA (1988) and Nissen et al. (2005) for recorded materials. The U87A P48 type Neuman microphone was used for the recording. The AKG cover/windshield was used over the microphone to prevent distortion, especially when plosive sounds were produced. The microphone was positioned approximately 15cm from the speaker at 0 degree azimuth. The microphone was connected to a sound mixer (Mackie - 32-8-2 8-Bus mixing console) in the control room. The recordings made were normalized to OdBHL levels. The recording of the words was preceded by a calibration pure tone signal which is according to the recent ISO 8253-3 specification for recorded materials. The calibration signal was weighted, using a frequency-modulated tone at 1 KHz,

which had a bandwidth of at least 1/3 octave (Nissen et al., 2005). The modulating signal was sinusoidal and had a repetition rate of 4-20Hz. The calibration tone lasted 60 seconds and was followed by an instruction read by a male Zulu first language speaker. The Zulu words were recorded in diverse, random sequences, on six different tracks on the CD. Fifty-eight words were recorded on each of the six tracks. The six tracks were randomizations of the same 58 words. The fifty-eight words were recorded with a 5 second pause between each word (Lyregaard, 1997). Each word was preceded by a carrier phrase "Yithi" which means "say" in Zulu. The Zulu First Language Speaker was advised to read the words in the natural tone that the words are spoken in accordance with the dialect used in the Durban-Pietermaritzburg region of KZN. The recordings were transferred from the mixer desk to the Carilon Audio system computer hard drive. The Wavelab software package was used to edit the words. The words were normalized to peak at 0dB and were adjusted once all recordings were done. The editing software allowed for the words to have the same RMS power as the 1000Hz calibration tone in an initial attempt to equate all the words. The use of a constant RMS allows for the output of the signal (the presentation of the words) to remain constant. This is a standard method of controlling any fluctuations in the signal and has been described in other studies such as Ramkissoon et al. (2002) and Nissen et al. (2005). The recordings were then produced onto a CD using the Nero Start Smart CD writer.

#### **The choice of speaker for the recording of the CD**

The speaker selected for the recording of the CD was a male Zulu first language university student who resided in the Durban-Pietermaritzburg region where Zulu is the main language spoken. Therefore, the dialect used in the recording was matched to that of the material selected and to the participants in this study. Recent research has indicated minimal gender influences on the recording of the material and the overall homogeneity of audibility of the words (Cambron, Wilson & Shanks, 1991; Nissen et al., 2005).

#### **Procedure**

##### **Procedure for the measurement of homogeneity of audibility**

The measurement of homogeneity of audibility using the recorded words occurred in two sessions i.e. session one involved testing of the participants for normal hearing and session two involved the measurement of homogeneity of audibility. Prior to any testing, the University of Kwa Zulu-Natal ethics committee approved the study. Informed consent was obtained from each participant before commencing with testing. The testing of the participants for normal hearing was conducted using standard audiometric procedures. These procedures included a detailed case history, otoscopic examination, immittance audiometry, air conduction testing for frequencies 125 Hz to 8KHZ. An isolated Industrial Acoustics Company (IAC) twin audiometric soundproof booth of double wall construc-

tion was used as the test environment for the pure tone testing and speech evaluation. These requirements were set by the American National Standards Institute [ANSI] (1977) to meet the noise level requirements. All equipment used in session one and two was calibrated prior to testing (June 2005). Appendix A reflects the pure tone averages for all participants.

Based on the results of the basic battery in session one, the participants were selected or excluded from the study. The excluded participants were referred for further diagnostic testing and appropriate management at the local clinic, hospital or university clinic. Thirty participants had normal hearing based on the results of the basic battery and were therefore included in session two i.e. for the measurement of homogeneity of audibility of the words.

The participants were seated in the test room while the researcher and two Zulu first language speaking scorers were in the control room. Each participant was instructed in Zulu by a Zulu first language speaker. The speaker was trained by the researcher in terms of the instructions and procedure to be followed. The instructions orientated the participants to the nature of the task. The instructions specified the participant's mode of response and indicated that the test material was speech. Instructions were given for participants to repeat the words that they heard verbally, including those words presented at the faintest listening levels. The verbal response is supported in the literature by Lyregaard (1997). The scoring of participant's responses was done independently by the researcher and the two Zulu first language speaking scorers. The researcher included this scoring method in order to improve the reliability of the scoring process. The 58 words (Appendix B) were captured onto a scoring form developed by the researcher. The participant's response was scored with a tick for an accurate, and a cross for an inaccurate response or if there was no response.

The researcher randomly selected the test ear because only one ear of each participant was tested for the homogeneity of audibility of the words. The participants had normal hearing in both ears and it did not matter which ear was selected for testing. The participants were instructed verbally by a Zulu first language speaker. Thereafter, the recorded material was presented to each participant. The recording began with the 1000Hz calibration tone, followed by a recorded instruction in Zulu. The instructions were routed through the GSI 61 audiometer using TDH-49 earphones. The intensity level of the instruction was at 30dBHL. At the onset of testing for homogeneity of the words, the attenuator dial was set at 0dBHL. At this level one recording of 58 words was played to the participant. The number of correct and incorrect words was recorded by each of the three scorers on the scoring form. The attenuator was raised in 5dB steps and a second recording was administered. A different recording was then played successively at each 5dB increment until 100%

identification was reached for all 58 words. The participants were given a 5 minute rest period after three recordings were played. The rest period was incorporated in order to improve co-operation from all participants and to eliminate fatigue. Furthermore, the participants were reinforced by the researcher through non-verbal reinforcements. All thirty participants were tested in the same way.

#### ***Procedure for analysis of homogeneity of audibility***

Fifty eight bisyllabic words were played back via a CD player to the participants. The participants were instructed to repeat the words heard. There were 3 scorers used to assess the responses of the participants in terms of the audibility of the words. The fifty eight words were presented at different sound intensity levels viz. at 0, 5, 10, 15 and 20 decibels. The Kappa test of agreement according to Agresti (1990) was used to assess the consistency of the scorers. The Kappa test was chosen since the responses were binary and it was a suitable test of agreement between two or more scorers. Results indicated a good agreement amongst all three scorers at 0.05 level of agreement. The good inter-rater reliability across all three scorers allowed the researcher to choose the scores of one of the Zulu first language scorers for application of the logistic regression equation.

The use of the logistic regression equation led to the analysis of homogeneity of audibility. Logistic regression was used since the data was binary (nominal) in nature. Appendix C describes the motivation and procedure used regarding the logistic equations, as stipulated by Agresti (1990). This was done using computerized statistical analysis i.e. SPSS version 11.6. A statistician at the Medical Research Council of South Africa (Durban) assisted with the statistical analysis. The average percentage correct identified by each participant for each word was calculated using the above equation. The performance intensity curve for the most homogenous words was identified. The steepness of the curve was calculated using a logistic regression plot. The raw scores for the 58 words were inserted into a logistic regression equation that was designed to calculate the percentage correct at each intensity level. Using the regression equation, the calculations for the range of threshold (50%) and the slope of the curve was ascertained for each word. This model provided the researcher with a method of estimating, at any given intensity, the performance of the words in terms of percentage. The most homogenous words were selected and the CD recording was edited. This consisted of 28 words. These words were then considered for the acoustic analysis of the prosodic features outlined in objective two of the study.

## OBJECTIVE 2: TO DESCRIBE THE PROSODIC FEATURES, SPECIFICALLY THE PITCH AND ENERGY CONTOURS OF THE MOST HOMOGENOUS WORDS SELECTED IN OBJECTIVE ONE

### Sample Material

The twenty-eight Zulu words identified as being most homogeneous were included as material for this objective of the study. The recorded words were used for the acoustic analysis.

### Apparatus and choice of tests

The Multispeech model 3700 by Kay-Elemetrics Corporation was the most suitable apparatus for the present investigation. It is a digitally driven system. The accompanying software programs enable the system to analyze the pitch contours of the signal, determine the fundamental frequency of the input signal and measure the spectrographic waveform patterns of the speech signal.

### Procedure

The input signal for the Multispeech 3700 analysis was via the CD recording. The Multispeech main program software was used. Four display window settings were created. Display A represented the digitized sound wave, display B was a wideband spectrographic analysis, and C and D represented the pitch contour and energy contours respectively. However, for the purposes of this study the pitch and energy contour settings were most relevant. Each of the 28 words was recorded on the CD and was analyzed acoustically. There was a separate pitch contour and energy contour for each word, which allowed for ease of analysis and comparison.

### Analysis

All 28 words were carefully analysed for the acoustic properties that were relevant for this study. The following parameters were used:

Pitch contours

Energy contours

The pitch contours and energy contours of each syllable within a word were analyzed via visual inspection. This entailed identifying the highest points for pitch and spectral energy for each syllable of the word, by placing the cursor on the highest point of the contour and reading the values. Thereafter, differences in the pitch and overall spectral energy for each syllable were determined using descriptive statistics. Data was tabulated. The difference values for pitch and energy were calculated as such values could be used to confirm the prosodic features of Zulu (E. Groenewald, Lecturer, University of Pretoria personal communication, October 2006). The use of pitch was particularly important as the prosodic features in a language are usually indicated by the rise and fall in the pitch of the word (Kent & Read, 1992). Furthermore, the similarities in the pitch and energy content of the words could confirm the homogeneity of the words acousti-

cally. The results of the study are outlined below.

## RESULTS

The results are presented according to each of the two objectives outlined in the Method.

### RESULTS FOR THE MEASUREMENT OF HOMOGENEITY OF AUDIBILITY

The data for the homogeneity of audibility was treated statistically using the logistic regression analysis technique. The estimates of the regression model are indicated in Table 1. All slopes and intercepts were significant at the 0.05 level of significance. The calculations of the threshold intensity and slope are also provided. These values are calculated using the equations outlined in Appendix C. Table 2 indicates the slope obtained at 50%, the slope at 20 to 80% and the estimated threshold obtained according to the regression model for each of the words.

Table 1. Estimates of the regression model

Word	$\alpha$	$\beta$	Word	$\alpha$	$\beta$
banga	-2.393	0.227	linda	-2.267	0.24
bheka	-1.351	0.179	loya	-3.017	0.237
bhema	-1.586	0.173	lunga	-3.856	0.262
chela	-3.014	0.307	minya	-3.662	0.237
cinga	-2.91	0.258	pheka	-1.897	0.162
dansa	-1.974	0.222	phonsa	-2.001	0.175
dinga	-2.7	0.133	qoba	-1.105	0.203
donsa	-2.165	0.234	sefa	-1.817	0.077
faka	-0.283	0.257	shada	-1.978	0.304
finya	-3.268	0.175	shaya	-0.433	0.186
geza	-1.85	0.228	thanda	-1.729	0.156
goba	-1.827	0.168	thatha	-2.689	0.242
gonda	-2.578	0.617	thela	-2.877	0.271
ggoka	0.104	0.167	thenga	-2.871	0.353
gxeka	-1.121	0.3	thola	-2.313	0.439
hlala	0.921	0.214	vala	-1.199	0.216
hleka	-1.205	0.19	veza	-3.942	0.273
hlenga	-2.347	0.226	vula	-2.794	0.207
hluba	-2.587	0.193	vuma	-2.527	0.195
holo	-2.524	0.178	washa	-1.109	0.383
jaha	-2.197	0.22	wina	-3.401	0.26
jeza	-2.94	0.322	xola	-2.262	0.321
khaba	-2.013	0.288	yanga	-1.908	0.261
khanya	-1.385	0.229	yeba	-2.566	0.261
kheta	-2.227	0.26	yeka	-2.061	0.316
kipha	-2.581	0.254	yenza	-2.222	0.214
landa	-1.377	0.178	yifa	-2.024	0.154
letha	-2.334	0.206	yona	-2.94	0.322
lima	-2.288	0.158	yosa	-3.489	0.2

The data recorded in Table 2 indicates that on average the estimated threshold intensity (softest level) at which all the words were heard was 9.66dBHL. On average the 50% correct recognition point occurred at about 9.66dBHL. The slope at 50% was an important statistic to calculate, since it represents the steepest slope on the logistic curve. This indicates that the rate of change

is greatest at this point. For a unit change in  $x$  (intensity level) we see the largest increase in the probability of audibility ( $p(x) = 1$ ). The measurement of homogeneity of the word list is dependent on those chosen to have a steep threshold slope.

Table 2. Summary of the slope at 50%; slope at 20 to 80% and the estimated threshold.

Word	Slope at 50% ( $0.5 \times 0.5^*b$ ) *100	Slope at 20 or 80%: ( $0.2 \times 0.8^*b$ )*100	(-) $a/b$ = threshold	Word	Slope at 50% ( $0.5 \times 0.5^*b$ ) *100	Slope at 20 or 80%: ( $0.2 \times 0.8^*b$ )*100	(-) $a/b$ = threshold
banga	5.675	<b>3.632</b>	10.542	linda	6	<b>3.84</b>	9.446
Bheka	4.475	<b>2.864</b>	7.547	loya	5.925	<b>3.792</b>	12.73
bhema	4.325	<b>2.768</b>	9.168	lunga	6.55	<b>4.192</b>	14.718
Chela	7.675	<b>4.912</b>	9.818	minya	5.925	<b>3.792</b>	15.451
cingga	6.45	<b>4.128</b>	11.279	pheka	4.05	<b>2.592</b>	11.71
dansa	5.55	<b>3.552</b>	8.892	phonsa	4.375	<b>2.8</b>	11.434
dinga	3.325	<b>2.128</b>	20.301	qoba	5.075	<b>3.248</b>	5.443
donsa	5.85	<b>3.744</b>	9.252	sefa	1.925	<b>1.232</b>	23.597
faka	6.425	<b>4.112</b>	1.101	shada	7.6	<b>4.864</b>	6.507
finya	4.375	<b>2.8</b>	18.674	shaya	4.65	<b>2.976</b>	2.328
geza	5.7	<b>3.648</b>	8.114	thanda	3.9	<b>2.496</b>	11.083
goba	4.2	<b>2.688</b>	10.875	thatha	6.05	<b>3.872</b>	11.112
gonda	15.425	<b>9.872</b>	4.178	thela	6.775	<b>4.336</b>	10.616
gqoka	4.175	<b>2.672</b>	-0.623	thenga	8.825	<b>5.648</b>	8.133
gxeka	7.5	<b>4.8</b>	3.737	thola	10.975	<b>7.024</b>	5.269
hlala	5.35	<b>3.424</b>	-4.304	vala	5.4	<b>3.456</b>	5.551
hleka	4.75	<b>3.04</b>	6.342	veza	6.825	<b>4.368</b>	14.44
hlenga	5.65	<b>3.616</b>	10.385	vula	5.175	<b>3.312</b>	13.498
hluba	4.825	<b>3.088</b>	13.404	vuma	4.875	<b>3.12</b>	12.959
hola	4.45	<b>2.848</b>	14.18	washa	9.575	<b>6.128</b>	2.896
jaha	5.5	<b>3.52</b>	9.986	wina	6.5	<b>4.16</b>	13.081
jeza	8.05	<b>5.152</b>	9.13	xola	8.025	<b>5.136</b>	7.047
khaba	7.2	<b>4.608</b>	6.99	yanga	6.525	<b>4.176</b>	7.31
khanya	5.725	<b>3.664</b>	6.048	yeba	6.525	<b>4.176</b>	9.831
kheta	6.5	<b>4.16</b>	8.565	yeka	7.9	<b>5.056</b>	6.522
khipa	6.35	<b>4.064</b>	10.161	yenza	5.35	<b>3.424</b>	10.383
landa	4.45	<b>2.848</b>	7.736	yifa	3.85	<b>2.464</b>	13.143
letha	5.15	<b>3.296</b>	11.33	yonia	8.05	<b>5.152</b>	9.13
lima	3.95	<b>2.528</b>	14.481	yosa	5	<b>3.2</b>	17.445

SUMMARY STATISTICS	Slope at 50%	Slope at 20 or 80%:	Threshold Intensity
Mean	5.986	<b>3.831</b>	9.657
Standard deviation	2.037	<b>1.304</b>	4.85
Maximum	15.425	<b>9.872</b>	23.597
Minimum	1.925	<b>1.232</b>	-4.304
Range	13.5	<b>8.64</b>	27.901

Figure 1. Performance intensity curves for Zulu words finally selected

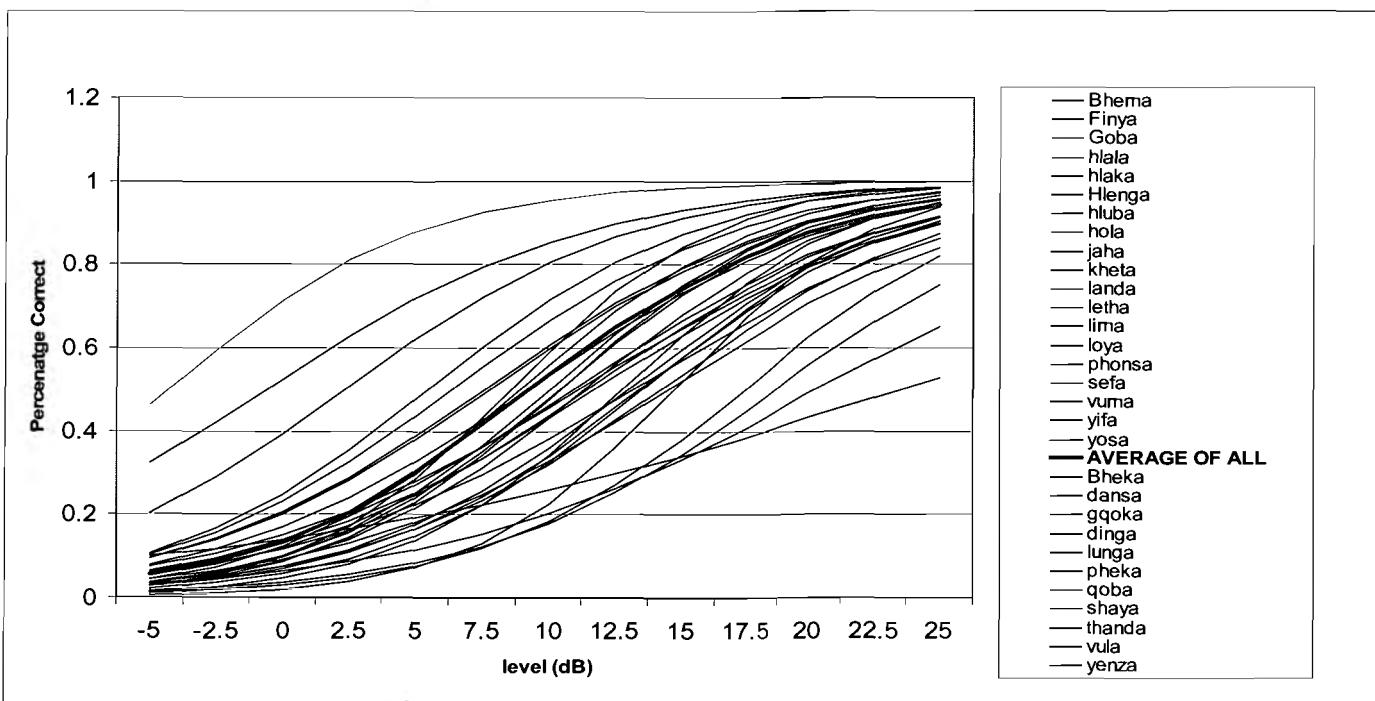
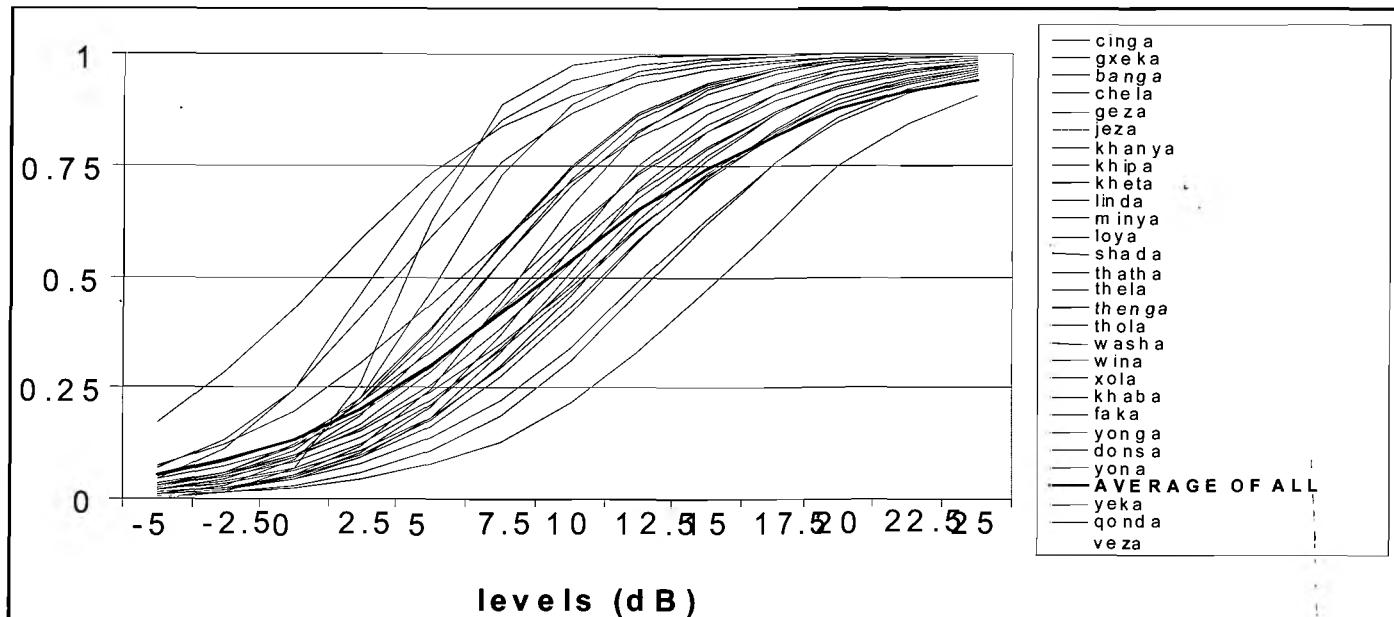


Figure 2. Performance intensity curves for Zulu words not selected



The mean slope at 50% was found to be 5.986%/dB ( $sd=2.037$ ). Therefore, the words that were above the mean slope were considered most homogenous. These words had the steepest slope and satisfied the criterion of homogeneity of audibility. Twenty eight words were therefore considered fairly homogenous on the basis of the above analysis. The performance intensity curve of the 28 words is described in Figure 1. A description of individual performance curves is available from the authors. Figure 2 illustrates the performance intensity curves for the words not selected. The mean slope of the words selected as indicated in Figure 1 was 5.9865%/dB. This figure also clearly indicates that the words selected had the steepest slope compared to the words not selected in Figure 2. Figure 2 indicates that the slope

at 50% for each of the words was more gradual. The gradual slope indicates that for these words to be identified correctly the words had to be presented at higher intensity levels compared to the words identified as being most homogenous

#### RESULTS EMERGING FROM THE ANALYSIS OF THE PROSODIC FEATURES OF THE WORDS

The twenty eight words selected in objective 1 were analyzed acoustically for their prosodic features. Table 3 indicates the results obtained for each of the twenty eight words. The majority of the words (89%) indicate a relatively lower pitch value in the first syllable compared to the second syllable. The pitch values in Table 3 also indicate a difference in pitch between syllable one and two. The magnitude of the difference value serves to con-

firm the change in pitch between syllable one and two. This confirms the presence of tonal patterns in the selected Zulu verbs i.e. low tone verbs.

the penultimate syllable, and therefore prominence is noted in the second syllable.

## DISCUSSION

No	Word	Pitch value	Pitch value	Difference	Energy	Energy	Difference
		Syllable 1	Syllable 2	in Pitch between syllables	value Syllable 1	value Syllable 2	in energy values between syllables
1	banga	107	113	6	67.83	73.9	10.1
2	gxeka	107	144	37	69.75	73	3.5
3	cinga	176	125	51	73.91	67.4	6.5
4	khanya	81	106	25	68.05	67.01	0.9
5	thela	106	135	29	60.3	68.8	8.5
6	khaba	108	132	24	69.12	75.4	6.2
7	kheta	104	125	21	55.1	60.9	5.8
8	jeza	79	152	73	69	69.18	0.2
9	donsa	139	93	46	69.9	69.31	0.6
10	khipa	85	121	36	66.38	72	5.6
11	chela	82	130	48	60.25	74.06	13.6
12	xola	109	131	22	66.64	60.3	1.3
13	yonga	88	104	16	70.28	69.26	1
14	yona	88	104	16	66.26	67.6	1.3
15	linda	87	98	11	65.91	66.08	0.8
16	veza	91	97	6	68.63	75.3	6.7
17	loya	106	127	21	72.06	63.6	8.5
18	thenga	151	141	10	60.2	46.1	14.2
19	wina	108	136	27	52.3	61.2	8.9
20	faka	81	106	25	68.63	75.8	7.2
21	washa	87	98	11	66.65	67.6	0.9
22	yeka	82	130	48	66.38	72	5.6
23	minya	105	151	46	49.2	56.9	7.7
24	thola	96	130	34	66.64	70.81	4.2
25	shada	85	121	36	70.26	74.81	4.6
25	shada	85	121	36	70.26	74.81	4.6
26	thata	85	127	42	68.63	74.81	6.2
27	qonda	107	139	32	69.75	72.81	3.1
28	geza	109	146	37	66.71	69.1	7.4

Table 3. Pitch and energy values: Syllable 1 and 2

The majority of the words (89%) show a similar pattern regarding pitch, and this serves to confirm the presence of prosodic patterns in Zulu. Eleven percent (3/28) did not indicate this pattern. This finding could have been influenced by the speaker placing artificial stress on the first syllable, similar to that when a spondee word is uttered. This limitation was expected as the speaker was a student of Audiology and the common use of spondee words in clinical practice may have influenced this production. The energy values of the vowels in each syllable also indicated a difference between syllable one and two. Generally, a higher energy value was noted in the second syllable of the words with the exception of seven words. The difference in the energy values obtained may also be due to the prosodic features of Zulu, whereby in isolation, low tone verbs may indicate lengthening of

Fifty eight bisyllabic low tone verbs were measured for homogeneity of audibility and were then further subjected to an acoustic analysis for the description of prosodic features. Twenty eight words were most homogenous, because they had the steepest performance curves. Statistically, words that were identified with a greater than 5.98%/dB slope, were considered to be most homogenous. The performance curve is used to illustrate how the words were heard at different intensity levels. The critical element of the performance curve is the steepness or the slope of the curve. The mean slope of these twenty eight words was 5.98%/dB. The mean slope, therefore, is comparable to that of other languages and this is indicated in Table 4 below.

Figure 1 also indicates variability in the curves generated. However, closer inspection of the twenty eight words selected (see Figure 1 in the results section) illustrates that the majority of the curves appear similar in shape with a distinct s-shaped curve. In contrast, however, the remaining 30 words were very different in shape (see Figure 2 in the results section). These

words indicated a more gradual slope than the words considered as most homogenous. The gradual slopes indicate that for these words to be identified correctly the words had to be presented at higher intensity levels compared to the words identified as being most homogenous.

The results obtained in the present study, when compared to studies of other languages indicate both similarities and differences. Table 4 provides a summary of the data obtained across different studies regarding the homogeneity of the word lists.

Researcher/s	Language	Type of stimuli and number of words evaluated	Scale of measurement	Mean Slope @ 50%
Present study	Zulu	Bisyllabic Verbs (28)	dBHL	5.98%/dB (6%/dB)
Hudgins et al. (1947)	English	Bisyllabic spondees (36)	dB	7.2%/dB
Young et al. (1982)	English	Bisyllabic spondees (15)	dB SPL	10%/dB
Cardenas & Marrero (1994)	Spanish	Trisyllables and four syllables (24)	dB	8%/dB
Christensen (1995), as cited in Nissen et al. (2005)	Spanish	Trisyllables (12)	dB	11.1%/dB
Greer (1997), as cited in Nissen et al. (2005)	Italian	Trisyllabic	dB	7.3%/dB
Nissen et al. (2005)	Mandarin	Trisyllabic (24)	dB	11.3%/dB
Nissen et al. (2008)	Taiwan Mandarin	Bisyllabic	dB	9.6/dB

Table 4. Overview of the mean slope values across studies.

Table 4 indicates that there are differences in the slope at 50% / dB across languages. The mean slope at 50% for the studies that utilized trisyllables compared to the bisyllables appear to be steeper e.g. 11.3%/dB in the Nissen et al. (2005) study compared to the 6%/dB in the present study. The steepness of the curve, however, is directly influenced by the type of stimuli used (Kruger & Kruger, 1997). In addition, the number of syllables in the stimuli influences the steepness of the slope. Trisyllabic or polysyllabic words are more redundant than bisyllabic words. Redundancy in speech audiometry is defined as "the abundance of information available to the listener due to substantial content of the speech material and the capacity of information inherent in richly innervated pathways of the central system" (Stach, 2003 p. 227). Thus, the abundance of information present within the speech signals of trisyllables or polysyllables allows for the words to be heard at softer levels, contributing to a steeper slope (Stach, 1998).

Furthermore, while the concept of redundancy is not highlighted as a criterion in the selection of SRT materials, Stach (1998) warns that the more redundant the stimuli is, the more immune the signal is to detecting a hearing loss. Therefore, the current researchers maintain the argument in support of using less redundant stimuli in the present study, i.e. bisyllabic Zulu words, as they may be more sensitive than trisyllabic or polysyllabic words in determining SRT.

Notwithstanding the differences obtained in the present study compared to other languages, the aim of determining which of the 58 preselected Zulu words met the criterion of homogeneity of audibility was realized in this investigation. An accurate replica of scores for the slope at 50% between and among studies is not possible, as languages used are different and the methodologies followed were different in determining the slope at 50%.

Another explanation for the more gradual slopes obtained in this study compared to the earlier studies conducted in English could be related to the differences in the linguistic structure of the languages. The use of spondees with equal stress explains the ease of identification at low intensity levels. Brandy (2002) confirms this by stating that the performance intensity function for spondaic words is very steep. The average level for 100% correct recognition to be obtained would usually occur at 27.5dB SPL (7.5dBHL). This is supported by earlier statements made by Egan (1948) that spondees have the highest homogeneity of audibility in comparison to other stimuli. It has also influenced subsequent researchers i.e. Ashoor & Proshazka, 1985; Plant, 1990; Harris et al. 2001; Nissen et al. 2005 to confirm homogeneity of audibility in their studies by drawing comparisons that are very closely related to the spondee words in English.

The criterion to assess the steepness of the slope for languages that do not have spondees e.g. Zulu, may however need to be altered. This is based on the fact that the Zulu words are different in stress pattern and structure to those of English, so the reference data used to evaluate the steepness of the slope should also be different.

There is a strong possibility that the normal hearing Zulu first language speakers in the present study may have had difficulty identifying the bisyllabic words in Zulu due to the lack of prominence of the first syllable, but required the intensity of the signal to be slightly higher for 100% recognition. This could, therefore, serve as a possible explanation for the mean threshold intensity for 100% correct recognition in the present study being 9dBHL.

In addition, the morphological structure of the bisyllabic verbs usually has a /-a/ bound to the verb root (Rycroft & Ngobo, 1979). The presence of the /-a/ sound in the second syllable together with the prominent high tone may account for the poor recognition of the first syllable. Thus, under difficult listening conditions, recognition of the first syllable may have been compromised, because of the lack of similar acoustic and morphological cues in the first syllable compared to the second syllable (Chetty, 1990).

Scrutiny of the performance intensity curves for the Zulu words indicates that the words that had the click sound in the first syllable e.g. /qonda/ and /gxeke/ had the steepest slopes at 50% i.e. 15.425%/dB and 7.5 %/dB respectively. These click sounds carry a higher spectral energy content acoustically, and thus, could account for the greater recognition of these words at the lower intensity levels. However, the words such as /finya; dinga; lima/

appeared to have had more gradual slopes at 50%. These words may be regarded as softer due to the fricatives and lower intensity vowel sound /i/ being used. The variation in the acoustical energy content of the consonants and vowels of Zulu may therefore have implications for the steepness of the slope. Chetty (1990) briefly alludes to the influence of the acoustic properties of vowels and consonants on the steepness of the curve. Within the present study, it would appear that the acoustic properties of the vowel and phoneme combinations could have contributed to the slope patterns observed. Apart from Chetty (1990), previous researchers have not referred to the influence of the energy content on the steepness of the slope. However, the present researchers are of the opinion that for a language like Zulu, the argument is relevant and worthy of noting. This point introduces a new lens through which homogeneity with respect to audibility is viewed in languages where spondees are not available. Further, qualitative comments made by participants indicated that they often heard the second syllable at the lower intensities, but had difficulty with hearing the first syllable. This reinforces the possible influence of prosody on the steepness of the slope.

Apart from the possible lower acoustic energy of the first syllable, the actual recording of the materials could have also affected the performance of the listeners. Nissen et al. (2005) explain the importance of high quality digital recording in the development of materials for SRT. While every attempt was made to adhere to this recommendation in the present study, the recording of the words in terms of VU meter peaking were +/-2dB across the words. Manual adjustments had to be introduced to the VU settings between the calibration tone and the words. Furthermore, the six randomizations were individually recorded due to a lack of software to help to randomize the recordings automatically. These limitations in recording, could have possibly affected the performance of the listeners with regard to the recognition of the words (refer to Panday (2006) for further explanation about the recording process in this investigation).

Despite the resource constraints and time limitations of this study, all data obtained were verified at several levels of analysis and therefore, remain reliable. Perhaps an implication of this would be to re-record the most homogenous words after introducing an intensity adjustment digitally. This has been done historically and more recently by Nissen et al. (2005): the variability of the curves is reduced and the steepness of the curves is improved. This could also account for the reason why the slopes obtained in the Nissen et al. (2005) study were similar to the data obtained in the original English studies.

From a methodological perspective, homogeneity of audibility has been assessed using different mathematical, statistical and procedural methods. Young et al. (1982) have raised the historical discrepancies noted in the English word lists regarding homogeneity of audibility. These researchers articulated concerns re-

garding the definition of homogeneity. Earlier studies (Bowling & Elpern, 1961; Curry & Cox, 1966; Beattie et al. 1975) reported on the intensity level at which words were first identified correctly, and made no provisions in their analysis if the word was missed in subsequent presentations (Young et al., 1982). In addition, the rate at which words were intelligible was not specified in the earlier studies. The present study considered these earlier limitations and utilized a more stringent mathematical model to calculate accurately the threshold of intensity and the slope of the psychometric function, taking into account the performance of participants across all lists. In addition, the rate at which words became intelligible can be easily calculated using this model. The model has also been described by Nissen et al.(2005).

While the present study followed a 5dB change in the intensity levels, the studies discussed above, including the Nissen et al. (2005) study used a 2dB intensity change. The use of the 5dB intensity change was adopted so as to maintain a consistent pattern with the manner in which SRT is usually established (Gelfand, 1997). The difference in intensity change, however, could account for the more gradual slope in the present study. The performance intensity curve with 2dB levels would have more reference points for calculating percentage correct responses.

The steepness of these curves would certainly improve the overall slope and shape of the curve. Future studies should consider decreasing the increments between the different intensity levels, thus ensuring that many more levels are assessed.

Another issue to be considered is the learning effect. Young et al. (1982) articulated concerns regarding the learning effect that could unfold if the same words are presented with several randomizations at the many reference levels discussed above. While the use of randomization is considered to eliminate the learning effect, one could also argue that with 13 presentations of the same words to a participant, there is a strong possibility that the participants learn the words and their performance would therefore appear better. This could also account for the excellent performance noted in the Nissen et al. (2005) study. Given the above, careful consideration must be afforded to the choice of intensity level changes and the number of randomizations selected. While the possibility of improving the performance curves exists, the potential impact of 'the learning effect' requires further investigation.

In summary, it is clear that selecting homogenous words in any language presents a challenge. Several factors can influence and affect the outcome. These may well include the linguistic, methodological and procedural variables. These challenges reinforce the need for the adaptation of the original criteria for SRT, at many levels, when materials are developed and selected for SRT testing. The reference data for evaluating homogeneity of audibility in English simply cannot be applied stringently to that of Zulu

or any other language. The concept of homogeneity of audibility is achievable in any language, but the criterion used to select the most homogenous words must suit that language.

This study described the selected words acoustically via an acoustic analysis system. The information obtained for the pitch contours is relevant to the present study. The majority of the words analyzed (25/28), indicated a similar pitch contour i.e. there was a difference in the pitch between the first and second syllable. Syllable one was consistently lower in pitch than syllable two. These findings follow the prosodic patterns of the Zulu words selected i.e. low tone verbs. Similarly the energy contours of the vowel sound in each word indicated differences between the two syllables. However, the patterns for 3/28 words indicated that the first syllable was higher in pitch than the second syllable. This finding could have been influenced by the speaker placing artificial stress on the first syllable, similar to a spondee word. A possible recommendation for future research would be to ensure that several recordings of the words are made and judged by listeners of the language before the final recording is accepted. This is supported by Nissen et al. (2005).

Nevertheless, the overall similarity between the pitch contours and energy contours for the majority of the words confirm that not only are the words homogenous in terms of linguistic and audiological variables, but most of the words were acoustically homogenous too. In addition, the use of the acoustic analysis to confirm the prosodic features in the language may be seen as an additional method for selecting words for SRT testing. This approach was clearly supported by Plant (1990), whereby spectrographic analysis supported the stress patterns of trisyllabic words in Tiwi. The application of acoustic analysis in the development of speech materials is in its infancy. However, the findings of the present study support the inclusion of acoustic analysis together with measurement of homogeneity of audibility as suitable 'tools' to be used in the selection of words for SRT testing.

## CONCLUSION

The findings in this investigation have important implications, both clinically and in terms of future research in the field. The Constitution of the Republic of South Africa (Act 200 of 1993), as amended by Act 109 of 1996, clearly confirms the importance of the eleven official languages of the country and the linguistic rights of individuals. In this context, the present investigation could be seen as a step towards the selection of SRT materials that are appropriate, relevant and scientifically designed to cater for the audiological needs of Zulu first language speakers in KZN. Similar research needs to be considered for the other official languages in South Africa, apart from English.

In terms of the diagnostic significance, it is clear that the measurement of SRT is fundamental to the initial audiological process. However, the value of the test is lost when the materials used are

not appropriately selected. In fact, the danger of over and under-diagnosis is evident. Thus, research of this nature provides the necessary steps towards improving the service delivery for Zulu first language clients in KZN. However, while the selection of an appropriate SRT word list was achieved in this study, the data obtained is limited to the adult normative population. Hence, there is a need for future research in terms of the wider clinical population.

## REFERENCES

- Agresti, A. (1990). *Categorical data analysis*. New York, USA: John Wiley Sons.
- American National Standards Institute (1977). *Standard criteria for permissibility of ambient noise during audiometric testing (ANSI S3-1977)* New York.
- American National Standards Institute (1996). *American National Specification for audiometers (ANSI S3.6-1996)*. New York.
- American Speech-Language-Hearing Association (1988). Guidelines for determining the threshold for speech. *ASHA*, 3, 85-88.
- Ashoor, A. A., & Prochazka, T. (1985). Saudi Arabic speech audiometry in children. *British Journal of Audiology*, 19, 229-238.
- Balkisson, Y. (2001). *Speech discrimination testing: Towards the development of wordlists for Zulu-first language speaking children*. Unpublished undergraduate research report. University of Durban-Westville, Durban, South Africa.
- Beattie, R. C., Svihovee, B. C., & Edgerton, B. J. (1975). Relative intelligibility of the CID spondees as presented via monitored live-voice. *Journal of Speech and Hearing Disorders*, 40, 84-91.
- Bess, F. H. & Humes, F. (2003). *Audiology The Fundamentals*. USA: Lippincott Williams & Wilkins.
- Borg, E., Wilson, M., & Samuelson, E. (1998). Towards an ecological audiology: Stereophonic listening chamber and acoustic environmental tests. *Scandinavian Audiology*, 27, 195-206.
- Bowling, L. S., & Elpern, B. S. (1961). Relative intelligibility of items on C.I.D. auditory test W-1. *Journal of Auditory Research*, 1, 152-157.
- Brandy, W. T. (2002). Speech audiometry. In J. Katz (Ed.), *Handbook of Clinical Audiology*. (pp. 96-110). Baltimore: Williams and Wilkins.
- Cambron, N. K., Wilson, R. H. & Shanks, J. E. (1991). Spondaic word detection and recognition functions for female and male speakers. *Ear & Hearing*, 12, 64-70.
- Cardenas, M. R. & Marrero, V. (1994). Speech audiometry in Spanish: new recorded lists. Proceeding of the 20<sup>th</sup> International congress of Audiology, Halifax.
- Chermack, G. D., & Musiek, E. F. (1997). *Central Auditory Processing Disorders New Perspectives*. London: Singular Publishers Group.
- Chetty, S. (1990). *An investigation into the effectiveness of two Zulu word lists in establishing speech reception thresholds in normal hearing bilingual (English/Zulu) subjects*. Unpublished undergraduate dissertation. University of Durban-Westville, Durban, South Africa.
- Cope, A.T. (1982). *A comprehensive course in the Zulu language*.

- Durban Howard College, University of Natal.
- Curry, E.T., & Cox, B.P. (1966). The relative intelligibility of items on CID Auditory Test W1. *Journal of Auditory Research*, 6, 419-424.
- Demody, P. & Lee, K. (1997). Speech tests at the National Acoustic Laboratories. In M. Martin (Ed.), *Speech Audiometry* (pp. 297-314). London: Whurr Publishers.
- Doke, C. M. (1930). *Textbook of Zulu Grammar*. RSA: Longman Green Company Ltd.
- Egan, J.P. (1948). Articulation testing methods. *Laryngoscope*, 58, 955-991.
- Fox, A. (2000). *Prosody*. New York: Oxford University Press.
- Gelfand, S. A. (1997). *Essentials of audiology*. New York: Thieme Medical Publishers.
- Grimes, J. (1992). *UCLA language materials Zulu language profile*. Retrieved November 02, 2004, from <http://www.lmp.ucla.edu/profiles/profz01.htm>
- Harris, R. W., Goffi, M. V. S., Pedalini, M. E. B., Gygi, M. A., & Merrill, A. (2001). Psychometrically equivalent Brazilian-Portuguese bisyllabic word recognition materials spoken by male and female talkers. *ProFono*, 13, 249-262.
- Harris, R.W., Kim, E.& Eggett, D.L. (2003). Psychometrically equivalent Korean bisyllabic words spoken by male and female talkers. *Korean Journal of Communication Disorders*, 8, 217-337.
- Harris, R. W., Nielson,W.S. McPherson,D.L., Skarzynski,H.& Eggett,D.L. (2004). Psychometrically equivalent Polish bisyllabic words spoken by male and female talkers. *AUDIOFOONOLOGIA*, 25, 1-15
- Hirsh, I.J., Silverman, S. R., Reynolds, E. G., Eldert, E., & Benson, R.W. (1952). Development of materials for speech audiometry. *Journal of Speech and Hearing Disorders*, 17, 332-337.
- Hudgins, C.V., Hawkins, J. E., Karlin, J. E., & Stevens, S. S. (1947). The development of recorded auditory tests for measuring hearing loss for speech. *Laryngoscope*, 57, 57-89.
- International Organization for Standardization (ISO) 8253-3 (1996). *Acoustic- audiometric test methods. Part 3: Speech Audiometry*. Geneva.
- Jacobson, M.& Trail, A. (1986). Assessment of speech intelligibility in South eastern Bantu languages: Critical considerations. *SAJCD*, 33, 15-28.
- Jerger, J. (1970). Clinical experience with impedance audiometry. *Archives of otolaryngology*, 92, 311-324.
- Kent, R. D.,& Read, C. (1992). *The acoustic analysis of speech*. London: Singular Publishing Group.
- Knight, J. (1997). Speech tests in some languages other than English. In M. Martin (Ed.), *Speech Audiometry*: (pp 315-324). London: Whurr Publishers.
- Konkle, D. F., & Rintlemen, W. F. (1985). *Principles of speech audiometry*. Baltimore: University Park Press.
- Kruger, B. & Kruger, F. M. (1997). Speech audiometry in USA. In M. Martin (Ed.), *Speech Audiometry* (pp. 233-277). London: Whurr Publishers.
- Lau, C. C., & So, K. W. (1988)! Material for Cantonese speech audiometry constructed by appropriate phonetic principles.
- British Journal of Audiology, 22, 297-304.
- Leedy, P. D., & Ormrod, J. E. (2005). *Practical research*. New Jersey: Prentice Hall.
- Lyregaard, P. (1997). Towards a theory of Speech Audiometry Tests. In M. Martin (Ed.), *Speech Audiometry* (pp. 34-62). London: Whurr Publishers.
- Maxwell, D.L. & Satake, E. (2006). *Research and Statistical methods in Communication Sciences and Disorders*. Canada: Thomson Delmar Learning.
- Nissen, S.L., Harris, R. W., Jennings, L. J., Eggett, D. L., & Buck, H. (2005). Psychometrically equivalent trisyllabic words for speech reception threshold testing in Mandarin. *International Journal of Audiology*, 44, 391-399.
- Nissen, S.L., Harris, R.W & Dukes, A. (2008). Word Recognition Materials for Native Speakers of Taiwan Mandarin. *American Journal of Audiology*, 17, 68-79.
- Panday, S (2006). The Development of a Speech Reception Threshold Test for Zulu First Language Speakers in Kwa-Zulu Natal. Masters thesis. University of Kwa-Zulu-Natal. Durban, South Africa.
- Panday, S.; Kathard, H.; Pillay, M.; & Govender, C (2007) The development of a Zulu Speech Reception Threshold Test for Zulu First Language Speakers in Kwa Zulu-Natal. *The South African Journal of Communication Disorders*, 54, 111-122.
- Plant, G. (1990). The Development of speech tests in Aboriginal languages. *Australian Journal of Audiology*, 13, 30-40.
- Punch, J. L., & Howard, M.T. (1985). Spondee recognition threshold as a function of set size. *Journal of Speech, Language & Hearing Disorders*, 50,120-125.
- Ramkisson, I. (2000). *Speech reception thresholds for non-native speakers of English: Digit pairs vs spondee words*. Unpublished doctoral dissertation, University of Illinois, Urbana Champaign, Illinois.
- Ramkisson, I., Proctor, A., Lansing, C. & Bilger, R. C. (2002). Digit speech recognition threshold (SRT) for non-native speakers of English. *American Journal of Audiology*, 11, 23-28.
- Robinson, D. D., & Koenings, M. J. (1979). A comparison of procedures and materials for speech reception thresholds. *Journal of American Audiology Society*, 4, 227-230.
- Roeser,J. R., Valente, M., & Horsford-Dunn, H. (2000). *Audiology Diagnosis*. USA: Thieme Medical Publishers.
- Rycroft, D. K., & Ngcobo, A. B. (1979). *Say it in Zulu*. Pietermarizburg:University of Natal Press.
- Schill, S. (1985). Threshold for speech. In J. Katz (Ed.), *Handbook of Clinical Audiology*. Baltimore: Williams & Wilkins.
- Silman, S., & Silverman, C. A. (1991). *Audiology diagnosis. Principles and practice*. New York: Academic press. I
- Stach, B. A. (1998). *Clinical audiology: An introduction*. San Diego: Singular Publishing.
- Stach, B. A. (2003). *Comprehensive Dictionary of Audiology Illustrated*. Canada: Library of Congress.
- Wilson, R.H., Zizz, C.A., Shanks, J. E., & Cuasey, G. D. (1990). Normative data in quiet broadband noise, and competing message for Northwestern University auditory test no.6 by a female speaker. *Journal of Speech Hearing Disorders*, 55, 771-778.
- Young, J.R., Dudley, B., & Gunter, M. B. (1982). Thresholds and psychometric functions of the individual spondaic words. *Journal of Speech and Hearing Research*, 25, 586-593.

## Appendix A

## Pure Tone Average scores

N	PTA(right)	PTA(left)
1	.00	.00
2	5.00	.00
3	13.30.	8.30
4	.00	.00
5	6.67	6.67
6	.00	.00
7	.00	.00
8	.00	-5.00
9	.00	.00
10	.00	-5.00
11	15.00	.00
12	.00	10.00
13	5.00	5.00
14	6.67	3.33
15	.00	.00
16	.00	.00
17	5.00	6.67
18	.00	.00
19	.00	.00
20	.00	.00
21	.00	.00
22	3.33	.00
23	5.00	5.00
24	3.33	5.00
25	13.33.	11.67
26	5.00	5.00
27	5.00	5.00
28	5.00	5.00
29	0.00	0.00
30	5.00	5.00

## Appendix B: Fifty eight words considered for assessment

banga	hlala	loya	veza
bheka	hleka	lunga	vula
bhema	hlenga	minya	vuma
chela	hluba	pheka	washa
cinga	hola	phonsa	wina
dansa	jaha	qoba	xola
dinga	jeza	sefa	yanga
donsa	khaba	shada	yeba
faka	khanya	shaya	yeka
finya	kheta	thanda	yenxa
geza	khipa	thatha	yifa
goba	landa	thela	yona
gonda	letha	thenga	yosa
gqoka	lima	thola	
gxeka	linda	vala	

## Appendix C

## Motivation for logistic regression

Logistic Regression:

Some basic background on the logistic model:

Since the binomial distribution is used, we might expect that there will be a relationship between logistic regression and chi-square analysis. It turns out that the 2x2 contingency analyses with chi-square is really just a special case of logistic regression. With chi-square contingency analysis, the independent variable is dichotomous and the dependent variable is dichotomous. Logistic analysis does not restrict the independent variable to be dichotomous.

Generally when one uses binary data the linear regression is not an appropriate model at all, in fact the data usually takes on an S-shaped curve. Our statistical effort to transform the S-shaped curve to a linear one requires us to employ the logistic distribution.

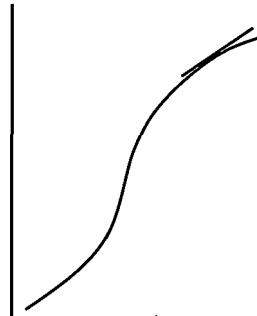


Figure: 1. S-shaped linear approximation to logistic regression curve.

For a binary response Y and a quantitative explanatory variable X, we allow  $p(x)$  to show the probability of a success when X takes on any of its values – for example in our data let X represent the sound intensity levels and Y denote the success or failure (audible or not respectively.) Then when  $x = 10$ , mathematically we write,

$$p(x) = \text{Prob}(y|x).$$

Read, the probability of y given x.

We also define the odds of success as:

$$\text{odds} = \frac{p(x)}{1 - p(x)}$$

The log of the odds is known as the logit.

The simple logistic model is of the form, (for only one explanatory variable, x):

$$\text{logit}[p(x)] = \log\left(\frac{p(x)}{1 - p(x)}\right) = \alpha + \beta x \quad (1)$$

This transformation, namely the logit (log of the odds), yields a linear function of the explanatory variables x. Estimation of the

parameters of interest in the logistic regression is done similarly to linear regression once the transformation has been applied. That is, we will have estimates for  $\alpha$  and  $\beta$  which represent the intercept and slope respectively (These are generally calculated by some software package). Solving for  $p(x)$  in equation (1) above gives,

$$p(x) = \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}}$$

(2)

Therefore,  $p$  is the proportion correct at any given intensity level. By inserting the intercept and slope into equation (2) we can predict the percentage correct at any given intensity level.

Any line drawn tangent to the curve in figure 1, has a slope equal to  $\beta * p(x) * [1 - p(x)]$ . The steepest slope occurs at  $p(x) = 0.5$ . The  $x$  value at this slope is usually called the threshold, (intensity required for 50% intelligibility). A simple equation to calculate threshold is given by

$$x = \frac{\log(p/1-p) - \alpha}{\beta}$$

(3)

At  $p = 0.5$ ,

$$x = \frac{-\alpha}{\beta}$$

(4)

#### Example of Calculation:

Estimates of  $\alpha$  and  $\beta$  are obtained from software package being used (here SPSS). Let us consider the word Geza at sound intensity level  $x = 5$ .

GEZA

$$p(5) = \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}} = \frac{e^{-1.85+0.228 \cdot 5}}{1 + e^{-1.85+0.228 \cdot 5}} = \frac{0.4916442}{1.4916442} = 0.32959$$

$$\begin{array}{cc} \alpha & \beta \\ -1.85 & 0.228 \end{array}$$

Thus the logistic model is of the form

And  $p(5)$  is

Thus we have the proportion correct at the 5dB level. For the 10dB level substitute  $x=10$  in the above. For the 15dB level substitute  $x=15$  in the above, and so on. Until we arrive at a list like the one below.

Word	$\alpha$	$\beta$	dB level (x)	$\exp(\alpha + \beta x)$	P(x)
geza	-1.85	0.228	-5	744	0.04788
geza	-1.85	0.228	-2.5	162	0.08166
geza	-1.85	0.228	0	717	0.135873
geza	-1.85	0.228	2.5	73	0.21755
geza	-1.85	0.228	5	42	0.329599
geza	-1.85	0.228	7.5	824	0.465057
geza	-1.85	0.228	10	752	0.605874
geza	-1.85	0.228	12.5	183	0.731059
geza	-1.85	0.228	15	819	0.827784
geza	-1.85	0.228	17.5	763	0.894731
geza	-1.85	0.228	20	755	0.937614
geza	-1.85	0.228	22.5	727	0.963736
geza	-1.85	0.228	25	632	0.979164

From (4) above we can work out the threshold value

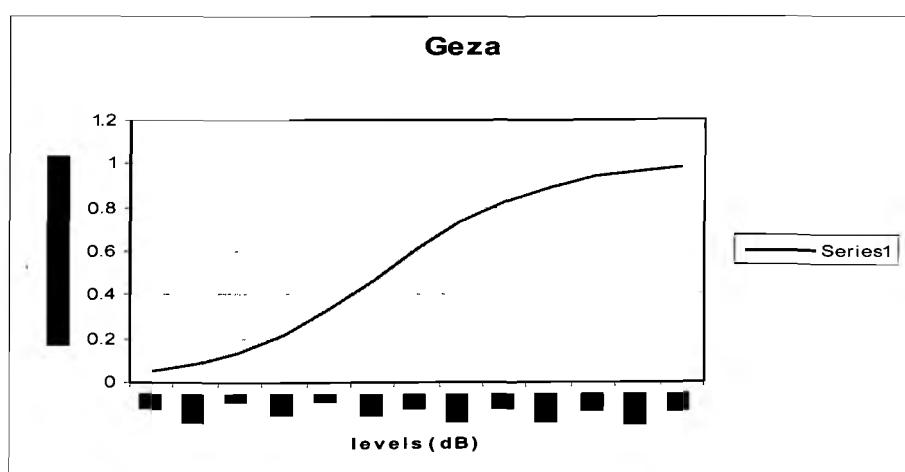
$$x = \frac{-\alpha}{\beta} = \frac{-(-1.85)}{0.228} = 8.1140$$

The slope at 50% ( $p(x) = 0.5$ ) or at threshold ( $x=8.1140$ ) is given by the formula:

$$p(x)[1 - p(x)]\beta 100 = 0.5(1 - 0.5)(0.228)100 = 5.7$$

Therefore the level at which audibility has a 50% chance occurs at 8.114 dB for the word Geza.

The plot of  $p(x)$  vs dB level ( $x$ ) gives the logistic curve below,



Calculations for the other words continue in a similar way.