# Clinical validation study

**Comparing results from hearSpeech to determine the patient's most comfortable level (MCL) and uncomfortable loudness level (UCL) for hearing aid fitting testing with traditional audiometer results to establish test accuracy, fail rate and test duration.**

## 5.1 INTRODUCTION

The purpose of this study was to compare the most comfortable level (MCL) and uncomfortable loudness level (UCL) determined by hearSpeech and a traditional audiometry testing in terms of test-retest reliability, correspondence and test duration.

hearSpeech includes uncomfortable (UCL) and most comfortable loudness (MCL) level testing and a speech-in-noise test using digits presented in noise. The UCL and MCL test component consists of three-digits presented at different intensities. The UCL and MCL loudness test is a subjective test and is determined based on the input responses. The speech-in-noise test component uses numbers in antiphasic and in-phase signals.

## 5.2 STUDY OBJECTIVES

To compare the UCL and MCL determined by hearSpeech and traditional audiometry in terms of test-retest reliability, correspondence and test duration.

## 5.3 METHOD

A sample of 20 subjects were recruited.

A within-subject design was used to determine the UCL and MCL results with two test methods (Method 1: Traditional and Method 2: hearSpeech) with a counterbalanced design.
The same pure-tone test, hearTest, was used to determine each test subject's pure-tone average (PTA) per ear as input criteria for each of the methods used for testing.

### *Method 1: Traditional*

Method 1 used a traditional audiometer with headphones with 7 loudness classification categories (1=Very soft; 2=Soft; 3=Comfortable, but slightly soft; 4=Comfortable; 5=Comfortable, but slightly loud; 6=Loud, but ok; 7=Uncomfortably loud). A stimulus was presented, which was live speech of various loudness levels. The test facilitator presented an initial stimulus at PTA+10dB. The stimulus continued to increase with +5dB steps until a response of '7=Uncomfortably loud' was obtained. The stimulus decreased with -10dB steps until "5=comfortable but slightly loud" or below was obtained. The cycle of +5dB up and -10dB down was repeated until two ratings of "6= loud but OK" were obtained at the same level during 3 consecutive ascending presentations. The level was recorded as the UCL for that test iteration. The procedure was repeated for another test iteration. If the UCL for the second test iteration differed by more than 5dB from the first test iteration, a third test iteration was conducted. The higher of the two values within 5dB of each other was taken as the UCL. MCL was determined after UCL was established by using the intensities that were captured from the user responses indicating a '4=Comfortable' rating. The final estimate of MCL was taken as the highest intensity where the user classified the loudness level as '4=Comfortable' across the two or three test trials. The test ended when a UCL and, or MCL value was determined. The UCL and MCL was determined for both ears.

### *Method 2: hearSpeech*

Method 2 used the hearSpeech application on a mobile device with headphones with 4 loudness classification categories (1=Too soft; 2=Comfortable; 3=Loud, but ok; 4=Uncomfortably loud). Randomised digit triplet sets was presented as the stimulus. The initial stimulus intensity of the test trial was dependent on the results of hearTest. Where the hearTest result was PTA<=60, the initial stimulus was PTA+20dB. Where the hearTest result was PTA>60db, the initial stimulus was PTA+10dB. In the case where a hearTest PTA value was not available, the initial stimulus was 50dB. The stimulus for PTA>60dB or where a hearTest PTA value was not available, increases in increments of +5dB, while the presented stimulus, where PTA<=60dB, increases in increments of +10dB. The test subject was required to select a loudness category. The input from the test subject was captured. If the input of the test

subject was '1= Too soft; 2= Comfortable; 3= Loud, but ok', the stimulus where the PTA>60dB or PTA value not available, increased in increments of +5dB, while the increments for the rest increased with +10dB. For the first instance where the response was '4= Uncomfortably loud', the $UCL_{first}$ value was captured. The stimulus then decreased by -10dB and continuously decreased in increments of -10dB, until the test subject responded '1= Too soft', '2=Comfortable', or '3=Loud, but ok'. The stimulus was again increased in increments of +5dB, until the test subject responded '4=Uncomfortable loud' and the $UCL_{second}$ was captured. $UCL_{first}$ and $UCL_{second}$ were compared for the test trial and the lowest UCL was captured as the UCL for the test trial. The test was repeated to get two UCL values. If the UCL values differed by more than 5dB, a third test was conducted. The difference between $UCL_{third}$ and $UCL_{second}$, and the difference between $UCL_{third}$ and $UCL_{first}$ was calculated. If the difference between the UCL values differed by 5dB or less, the average of the two UCL values, where the difference was the lowest, was calculated as the UCL result. If all the UCL values differed by more than 5dB, the lowest UCL value was taken as the UCL result.

If the user did not respond by indicating that the tone was 'Uncomfortably loud' with the first test iteration, a second iteration was done. If, during the second iteration, a response of 'Uncomfortably loud' was captured, the intensity was recorded as the UCL for the test. If an 'uncomfortably loud' response was not captured during the second iteration, it was assumed that the test subject had an UCL at an intensity that exceeded the output range of the hearSpeech test (90+ dB).

If the user responded unexpectedly in a test iteration (ie. 'Soft' 'Loud, but ok', 'Soft'), a flag was triggered on the test and a third test iteration was required.

The intensity where the test subject responded '2= Comfortable' was captured as the MCL value. The MCL result was calculated as the highest iteration where the user responded as most comfortable for test trial 1 and test trial 2 (and test trial 3- if available). The test ended when a UCL and MCL value was determined. The UCL and MCL was determined for both ears.

### *Equipment*

Data collection was conducted in a sound booth at the University of Pretoria at the Department of Speech Language Pathology and Audiology. The booth was fitted with a chair with a backrest for the test subjects to be seated in during the test duration. The subjects were provided Samsung Tab A (10'') devices with Sennheiser HD300 calibrated headphones and conducted tests on a traditional audiometer.

### *Procedures*

The test facilitator ensured that the test subject was comfortably seated in front of a tablet. The test facilitator explained the test instructions to the test subject. The test facilitator placed the headphones on the ears of the test subject and requested the test subject to click on the screen each time a tone is heard.
A hearTest was conducted per test subject to determine the Pure Tone Average (PTA) value for each ear.

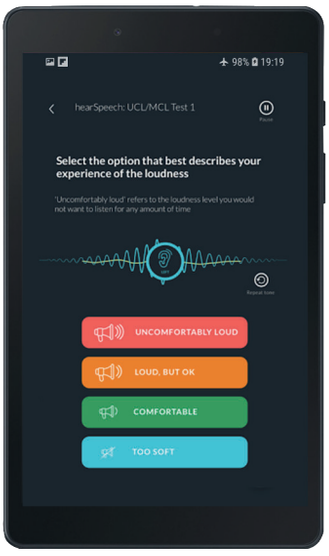
The UCL and MCL test is conducted in a traditional soundbooth setup, using an audiometer and headphones with 7 loudness classification categories (1=Very soft; 2=Soft; 3=Comfortable, but slightly soft; 4=Comfortable; 5=Comfortable, but slightly loud; 6=Loud, but ok; 7=Uncomfortably loud) and the test facilitator followed method 1 with input from the test subject.

The UCL and MCL test is conducted in a test setup using the mobile device, headphones and the hearSpeech application with 4 loudness classification categories (1=Too soft; 2=Comfortable; 3=Loud, but ok; 4=Uncomfortably loud). The test facilitator explains the test procedure of using the hearSpeech application to the test subject and the test subject conducts the test on their own using a Samsung Tab A device. The test subject follows method 2 without assistance from the test facilitator and chooses the relevant loudness category on the screen as the stimulus gets presented.

The UCL and MCL results were displayed on the result screen.

The facilitator alternated the test order to counterbalance the data collection. Test duration were determined by measuring the duration of each test Individually.

Figure 1. Loudness level options for selection using the hearSpeech application.




## 5.4 RESULTS

The result for UCL and MCL was collected and will be discussed in this section.

### 5.4.1 UCL RESULTS

#### *Test-Retest reliability*

The average UCL difference of test-retest results as obtained by the hearSpeech application was -1.8 dB with a standard deviation of 3.6 (Table 1). The hearSpeech UCL test-retest results were highly correlated (Figure 2). The average UCL difference of test-retest results as obtained by the traditional audiometer was 0.0dB with a standard deviation of 2.6 (Table 1). The traditional audiometry UCL test-retest results were highly correlated (Figure 3).

Table 1: Test-Retest reliability between UCL results for hearSpeech and traditional audiometer (n=39)

| | hearSpeech | Traditional audiometer |
|---|---|---|
| n | 39 | 39 |
| Average test UCL (SD) | 84.5 dB (9.8) | 83.0 dB (10.8) |
| Test UCL range | 60 to 95dB | 60 to 95 dB |
| Average retest UCL (SD) | 86.32dB (8.9) | 83.0 (10.7) |
| Retest UCL range | 60 to 95dB | 60 to 95 dB |
| Average difference (SD) | -1.8 dB (3.6) | 0.0 dB (2.6) |

Figure 2. Test-Retest reliability between UCL results from the hearSpeech application.

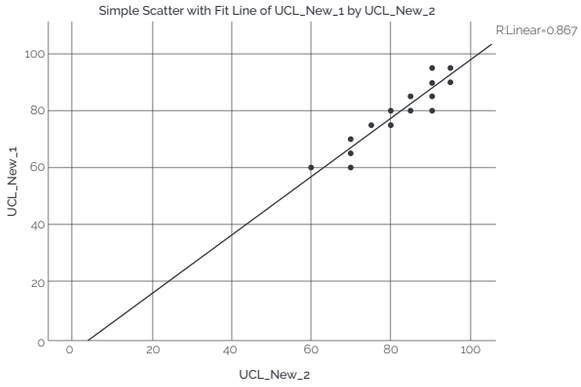
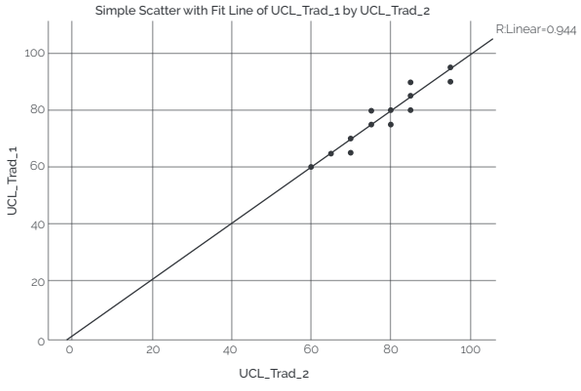



Figure 3. Test-Retest reliability between UCL results from the traditional audiometer.



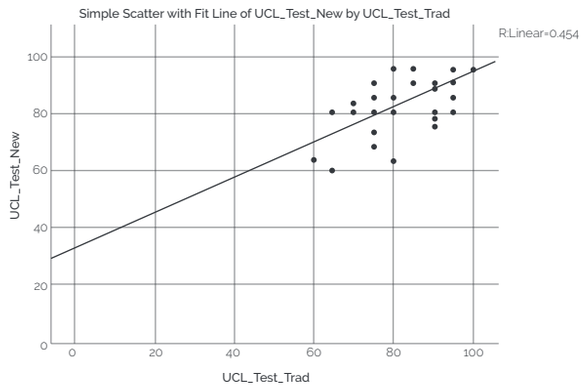

#### *UCL correspondence between hearSpeech and traditional audiometry*

The average difference between the hearSpeech and traditional audiometry UCL was -0.3dB with a standard deviation of 8.3 (Table 2). The UCL results were highly correlated between hearSpeech and the traditional audiometry (Figure 4).

Table 2: Comparison between UCL reliability for hearSpeech and traditional audiometer (n=39)

| | hearSpeech | Traditional audiometer | Difference |
|---|---|---|---|
| n | 39 | 39 | - |
| Average UCL result (SD) | 83.8 dB (9.8) | 83.6 dB (10.6) | -0.3 dB (8.3) |
| Test UCL Range | 60 to 95dB | 60 to 100dB | - |

Figure 4. Correlations between overall UCL test results obtained for hearSpeech and the traditional audiometer.




## 5.4.2 MCL RESULTS

### *Test-Retest reliability*

The average MCL difference of test-retest results as obtained by the hearSpeech application was -3.9dB with a standard deviation of 10.7 (Table 3). The hearSpeech MCL test-retest reliability was highly correlated (Figure 5). The average MCL difference of test-retest results as obtained by the traditional audiometer was -0.3dB with a standard deviation of 4.5 (Table 3). The traditional audiometry MCL test-retest results were highly correlated (Figure 6).

Table 3. Most comfortable listening level (MCL) comparison between hearSpeech and traditional audiometer (n=39)

| | hearSpeech | Traditional audiometer |
|---|---|---|
| n | 39 | 39 |
| Average test UCL (SD) | 66.2 dB (13.2) | 75.7 dB (11.3) |
| Test UCL range | 30 to 85dB | 50 to 95dB |
| Average retest UCL (SD) | 70.1 dB (10.6) | 75.9 dB (11.6) |
| Retest UCL range | 50 to 90dB | 55 to 95dB |
| Average difference (SD) | -3.9 dB (10.7) | -0.3 dB (4.5) |

Figure 5. Test-Retest reliability between MCL results obtained for hearSpeech application.

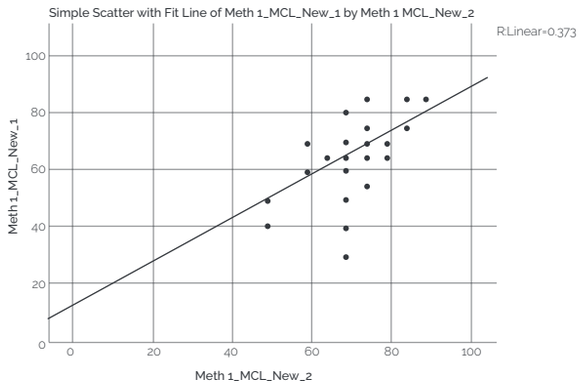
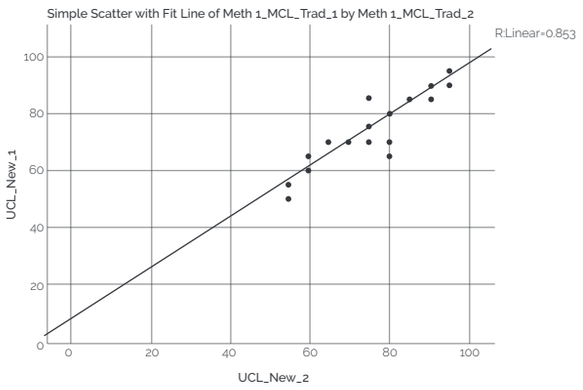



Figure 6. Test-Retest reliability between MCL results obtained for traditional audiometer.



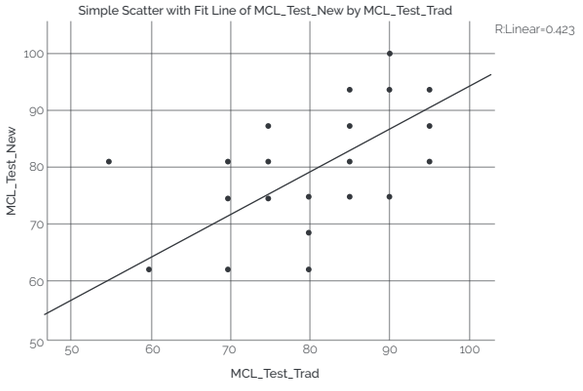

### *MCL correspondence between hearSpeech and traditional audiometry*

The average difference between the hearSpeech and traditional audiometry MCL was -5.4dB with a standard deviation of 9.1 (Table 4). The MCL results were highly correlated between hearSpeech and the traditional audiometer (Figure 7).

Table 4: Comparison of MCL reliability for hearSpeech and traditional audiometer (n=39)

| | hearSpeech | Traditional audiometer | Difference |
|---|---|---|---|
| n | 39 | 39 | - |
| Average UCL result (SD) | 71.7 dB (10.5) | 77 dB (11.2) | -5.4 dB (9.1) |
| Test UCL Range | 50 to 90 dB | 55 to 95 dB | - |

Figure 7. Correlations between overall MCL test results obtained for hearSpeech and the traditional audiometer.




## 5.4.3 TEST DURATION BETWEEN HEARSPEECH AND TRADITIONAL AUDIOMETER

The hearSpeech application is 50 seconds (30-seconds standard deviation) faster on average than the traditional audiometer (Table 2). This translates to a 49% improvement in test time using hearSpeech compared to a traditional audiometer.

Table 4. Time comparison when determining UCL and MCL from the hearSpeech application and traditional audiometer (n=39)

| | hearSpeech | Traditional audiometer |
|---|---|---|
| **n** | 39 | 39 |
| **Average test duration (SD)** | 52.1 sec (21.1) | 102.1 sec (24.2) |
| **Range** | 19 sec- 116 sec | 58.5 sec- 156 sec |

## 5.5 CONCLUSION

The hearSpeech UCL was reliable and accurately corresponded with traditional audiometry testing. hearSpeech UCL corresponded with 1 dB on average with traditional UCL audiometry testing and test-retest reliability was within 2 dB for both techniques. hearSpeech MCL was systematically lower than traditional audiometry MCL, but within acceptable clinically ranges for test-retest and correspondence. The hearSpeech test duration offered a 49% reduction in test time.