Developing an Articulation Test for Arabic-Speaking School-Age Children
Introduction
Arabic is a Semitic language spoken by more than 422 million people (Microsoft, Encarta encyclopedia®, 2004).

The classical Arabic language is utilized only on the academic and religious levels, and there are many colloquial Arabic dialects exhibited by Arabian which are extremely different in each Arabian country.

The consonant inventory of CEA consists of 27 consonants and 8 vowels (Ammar and Morsi, 2006).

CEA includes the primary emphatic phonemes /t/, /d/, /s/, /z/, which distinguish the Arabic language. These emphatic phonemes physiologically differ from their non-emphatic counterparts' /t/, /d/, /s/, /z/.
There are eight vowel phonemes in CEA. Five long vowels / i:, e:, a:, o:, u:/ and three short vowels / a, i, u/ (Ammar and Morsi 2006). Vowels do not occur in initial position and there are no diphthongs in CEA (Harrell 1957).

The need to test the children whom mother tongue is Arabic led to emergence of many articulation tests (Kotby et al., 1985; Amayreh and Dyson, 1998; Morsi, 2003 and Hamdan and Amayreh, 2007).
However, not all CEA consonants are represented in these tests; equally, not all the possible consonant' positions in the word are tested and none includes explicit vowel testing in their construction.

The aim of the work
The aim of this prospective pilot study was to develop an Arabic articulation test using familiar, culturally-based and visually transparent words in order to be used as a criterion for comparing phonemes of both normal and phonologically disordered Arabic-speaking children.

Subjects and methods
Mansoura Arabic Articulation Test (MAAT) passed into the following stages to be established:

- Stage of Designing.
- Stage of validation.
- Stage of test application.
- Reliability stage.

**Stage of Designing:**

A 106 pictures-naming test was designed to elicit spontaneous single word responses representing all possible initial, middle, final and double positions of CEA consonants and the eight CEA vowels.
Stage of validation:

Validity refers to the degree to which the test or other measuring device is truly measuring what it intends it to measure. Three judges, who are expert phoniatricians were asked to review the constructed test and respond to a questionnaire covering the following items:

Judges questionnaire:

- Chosen age group
- Suitability of the words to the tested consonants/vowel position
- Suitability of adding double consonant word to
- Suitability of adding basic vowels words to the test
- Allowing the child to pronounce the word after the examiner in cases of difficult words
- The use of cues to elicit the child response to pictures in stead of direct pronunciation of difficult words
Stage of test application:

The MAAT was presented to a randomly selected 100 Arabic speaking Egyptian normal children selected from the first and second grade kindergarten. They were 52 males and 48 females with their ages ranged between 42-70 months (X= 56.1 ± 7.96 months).

Stage of test application:

Each child was tested individually in a 15-20 minute session. A picture-naming was used to elicit the child's response. Children's responses were converted to a percent-correct score for sound utterance and picture identification.
Reliability stage:
✓ Reliability is synonymous with the consistency of a test. Test-Retest reliability refers to the test’s consistency among different administrations.
✓ MAAT reliability was examined using test-retest method and the same children were retested five weeks latter with the same examiner under same circumstances.

Statistical analysis:
✓ Frequency, mean and standard deviation were used to describe data.
✓ T-test was applied to compare two means and ANOVA was used to compare more than two means.
✓ Pearson correlation Coefficient was used to detect linear relationship between two quantitative variables. Spearman’s rho correlation coefficient was used for estimating test reliability.
**Results**

- Interpretation of scoring results of MAAT is presented in three sections:
  - Judgments' Validity results.
  - Results of test application.
  - Reliability results.
Judgments' Validity results: Comparison of the judges' opinions to the questionnaire items in % scores:

<table>
<thead>
<tr>
<th>Judges</th>
<th>Suitability of the tested word</th>
<th>Suitability of the representing pictures</th>
<th>Other Questionnaire items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>First</td>
<td>86.73 ± 0.34</td>
<td>97.96 ± 0.14</td>
<td>100 ± 0.00</td>
</tr>
<tr>
<td>Second</td>
<td>88.78 ± 0.32</td>
<td>96.94 ± 0.17</td>
<td>100 ± 0.00</td>
</tr>
<tr>
<td>Third</td>
<td>86.73 ± 0.34</td>
<td>95.92 ± 0.20</td>
<td>87.5 ± 0.35</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>87.42 ± 1.18</td>
<td>96.94 ± 1.02</td>
<td>95.83 ± 7.22</td>
</tr>
<tr>
<td>F</td>
<td>0.39</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>P</td>
<td>0.680</td>
<td>0.372</td>
<td>0.351</td>
</tr>
</tbody>
</table>

F=Repeated measures ANOVA, P > 0.05 (non-significant).

Results of test application.

Descriptive data for % correct scores for sounds' utterances and identified pictures after the first and second applications of MAAT:

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Middle</th>
<th>final</th>
<th>Double</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>Correct consonant utterance</td>
<td>1st</td>
<td>100 ± 0.0</td>
<td>100 ± 0.0</td>
<td>100 ± 0.0</td>
<td>97.64 ± 7.1</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>100 ± 0.0</td>
<td>100 ± 0.0</td>
<td>100 ± 0.0</td>
<td>99.04 ± 3.6</td>
</tr>
<tr>
<td>Identified pictures for consonants</td>
<td>1st</td>
<td>99.92 ± 0.4</td>
<td>97.64 ± 7.4</td>
<td>95.4 ± 17.3</td>
<td>83.1 ± 24.0</td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td>98.56 ± 4.0</td>
<td>98.96 ± 3.8</td>
<td>99.56 ± 2.2</td>
<td>94 ± 8.4</td>
</tr>
<tr>
<td>Correct Vowel utterance</td>
<td>1st</td>
<td></td>
<td></td>
<td>100 ± 0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td></td>
<td></td>
<td>100 ± 0.00</td>
<td></td>
</tr>
<tr>
<td>Picture identification for vowels</td>
<td>1st</td>
<td></td>
<td></td>
<td>96 ± 17.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2nd</td>
<td></td>
<td></td>
<td>100 ± 0.00</td>
<td></td>
</tr>
</tbody>
</table>
Results of test application.

Discriminant validity for the MAAT items:

<table>
<thead>
<tr>
<th></th>
<th>Group(1) &lt; 4 ys (17)</th>
<th>Group(2) 4-5 ys (44)</th>
<th>Group(3) &gt;5 ys (39)</th>
<th>Total</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td></td>
<td>0.00*</td>
</tr>
<tr>
<td>Number of Correct CU</td>
<td>97.35±0.70</td>
<td>97.77±0.52</td>
<td>97.76±0.27</td>
<td>97.76±0.51</td>
<td>8.37</td>
<td></td>
</tr>
<tr>
<td>Number of Correct VU</td>
<td>8 ±0.00</td>
<td>8±0.00</td>
<td>8±0.00</td>
<td>8±0.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Number of Identified pictures for C</td>
<td>97.29±0.69</td>
<td>97.77±0.52</td>
<td>97.92±0.27</td>
<td>97.75±0.52</td>
<td>10.4</td>
<td>0.00*</td>
</tr>
<tr>
<td>Number of Identified pictures for V</td>
<td>8 ±0.00</td>
<td>8±0.00</td>
<td>8±0.00</td>
<td>8±0.00</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

ANOVA test, * P < 0.001 (highly significant), CU=consonant utterance, VU=vowel utterance

Results of test application.

Content validity:

It is the extent to which the content of a test is representative of the conceptual domain that designed to cover. A statistically highly significant adequate correlation was found between correct word utterance and picture identification (r=0.91, P=0.00).
Reliability results.
Spearman’s correlation coefficient for reliability of MAAT according to age and gender:

<table>
<thead>
<tr>
<th>Age</th>
<th>Sex</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;5 years</td>
<td>&gt;5 years</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Correct word Utterance</td>
<td>0.78</td>
<td>0.80</td>
<td>0.70</td>
</tr>
<tr>
<td>Picture identification</td>
<td>0.70</td>
<td>0.46</td>
<td>0.58</td>
</tr>
</tbody>
</table>
MAAT was constructed to include all the CEA consonants in all possible positions in the word. Double consonant' position was included with the hypothesis to help in provocation of the subject’s articulatory break down. The CEA vowels was also included to be used in diagnosis of speech and language disorders that may include vowel distortions e.g. dysarthria.
The CEA vowels was also included to be used in diagnosis of speech and language disorders that may include vowel distortions e.g. dysarthria.

Statistics for judge validity for the MAAT revealed non significant differences among the judges' opinions with the observation that the percentages of agreement of suitability of the words representing the tested sounds and their representing pictures from the judges' point of view and the percentages of agreement to the other questionnaire items exceeded 85%.
The descriptive analysis of the children's accurate responses was not markedly changed for pictures replaced. This suggests that vocabulary rather than visual imagery is the important variable in picture ambiguity. Moreover, the statistically highly significant adequate correlation between correct word utterance and picture identification proved the content validity of MAAT.

The strong test-retest reliability for correct sound utterance and the moderate reliability for pictures identification proved the consistency of MAAT items.
Conclusion

- MAAT is a valid and reliable test that could be applied to collect phonetic inventory of Arabic speaking young children.
- Further studies are recommended to apply MAAT on patients with phonological and speech disorders.
Thank you