The Central Claim
Using a large corpus of natural data, it is possible to show “rhythm” differences statistically and graphically between Modern Standard Arabic (AR), American English (EN) and Chinese (Modern Standard Mandarin) (CH) using multiple speech timing pattern metrics.

Background Information
Typical examples — Syllable timing: Spanish, French, Chinese
Speech timing: English, German
Hamdi 2007 shows timing differences in Arabic dialects but does not include “Standard Arabic”

Neuer Work

Dockendorf et al. 2008 (this study)
AIM: To show that like English and Chinese, Arabic also has a “rhythm” continuum across multiple speakers in a large corpus of more natural speech.

50+ minutes of broadcast news speech (Voice of America)
46 Arabic News Reporters (those with affiliation to the broadcasting agency)
Replicated Benton et al. ’07 with Arabic data & interpolated with previous results

Our Source Data
Audio collected by Linguistic Data Consortium (LDC):
Chinese data sets from CCTV Beijing – (100 minutes, 21 Male, 29 Female spks).
English data from ABC and CNN news – (75 minutes, 27 M, 15 F spks).
Arabic data from Voice of America news – (50+ minutes, 24 M, 22 F).

Methods
Followed RNM, GL, & Benton et al. ’07
Phoneme durations calculated by speech recognition alignment of audio and transcription data.
Scripts to automate ΔC, %V, nPVI, pPVI, etc. calculations within sentences
Additionally divided and plotted speakers base on gender, and genre.
(For Arabic only used genre of news broadcasters at this point).
Disregarded any speaker whose contribution was less than 0.01% of total data for his or her language an artifact.

Definitions and Explanation of Terms
PVI: Pairwise Variability Index (nPVI: normalized, pPVI: raw)

Rhythm class: Used in the same sense as it in by those whose methods we followed — speech timing patterns not meterical rhythm.

Speech Rate: Assumption—nPVI normalizes for speech rate.

Utterance: Sentence-like unit of oral language as divided by a native speaker.

Genre: Could also be register, formal vs. informal, spoken vs. read aloud, etc.

Future Research for Arabic:
Add different Genre data (only tested news reporters) and categorize between Reporters and Interviewees.
Add other News data from other AR Dialects.
Add Conversational AR speech.

Table 1: Shows the numerical values (from the graph on the right) for each metrics for each language — VnPVI= Vocalic normalized PVI, CrPVI = Consonantal/Intervocalic Raw PVI

<table>
<thead>
<tr>
<th></th>
<th>%V</th>
<th>ΔC</th>
<th>VnPVI</th>
<th>CrPVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>35.12</td>
<td>49.13</td>
<td>39.32</td>
<td>78.55</td>
</tr>
<tr>
<td>Chinese</td>
<td>39.10</td>
<td>31.33</td>
<td>49.30</td>
<td>49.32</td>
</tr>
<tr>
<td>English</td>
<td>41.81</td>
<td>36.20</td>
<td>60.02</td>
<td>62.30</td>
</tr>
</tbody>
</table>

Results
SPSS ANOVA — significant difference between AR, CH, and EN all variables
Independent T-test — significant difference only for AR reporters on AC
Higher ΔC & CrPVI (even with lower %V & VnPVI) suggests AR is more Stress-Timed

Gender — No statistical difference within AR data set

Figure 2: Individual speaker diversity from all 3 languages CH, EN & AR (CH – boxes | EN – circles | AR – diamonds)

Figure 3: The division of news reporters by language, gender is shown by overlapping layers representing the boundaries of each group.

CONCLUSIONS
Analysis of naturally occurring data by multiple speakers from each language suggests that a inter-language “rhythm continuum” exists such that rhythm may be relative to the speaker and the context in which the language is used.
Support of a mean rhythm classification of Mod. Stand. Arabic based on the means of multiple speakers.
Phonetic procedures for determining rhythmic categories do extend to more naturally occurring Arabic data.
Like English and Chinese there appears to be a local continuum of speech rhythm/timing in Arabic.

A special thanks to our colleagues, administrators, professors, and those researchers at UTA and UTD who offered guidance and data without which this project would not have been possible.