

# Evaluation and Result

(Please open in Adobe Reader 9 or Higher to listen to the embedded audio files)

Language: English				
Target audio	Synthesized audio	Text sentence	Phoneme sequence	Cross correlation
		The Roman letter was used side by side with the Gothic	ði rəʊmən lətər wɒz juzd səd bəi səd wið ði gθɪk	0.6562
		Were induced to cut punches for a series of "old style" letters	wɜːr ɪndʒust tu kʌt pʌntʃɪz fər eɪ sərɪz ɒv əʊld stɔːl lətəz	0.66480
		It is obvious that legibility is the first thing to be aimed at in the forms of the letters	ɪt ɪz ɒbviəs ðæt ledʒɪbilətɪ ɪz ði fəst θɪŋ tu bi əimd æt m ði fəmz ɒv ði lətəz	0.5809

Table 1: Objective evaluation Of English synthesizer(LJSpeech dataset).

Language: English				
Target audio	Synthesized audio	Text sentence	Phoneme sequence	Cross correlation
		It fairly clubbed me into recognizing it	ɪt feəlɪ kləbd mi: mtu: rækəgnəɪzɪŋ ɪt	0.5938
		They were three hundred yards apart	ðeɪ wɜːr θri: hʌndrəd ja:dz əpə:t	0.6610
		It was a curious coincidence	ɪt wɒz eɪ kjuərɪəs kəʊmənsɪdəns	0.6924

Table 2: Objective evaluation of English synthesizer (CMUArctic male).

Language: English				
Target audio	Synthesized audio	Text sentence	Phoneme sequence	Cross correlation
		It fairly clubbed me into recognizing it	ɪt feəlɪ kləbd mi: mtu: rəkəgnaɪzɪŋ ɪt	0.7188
		They were three hundred yards apart	ðei wɜːr ðri: hʌndrəd ja:dz əpart	0.7090
		It was a curious coincidence	ɪt wɒz eɪ kjʊəriəs kœvənsɪdəns	0.5864

Table 3: Objective evaluation of English synthesizer (CMUArctic female).

Language: German				
Target audio	Synthesized audio	Text sentence	Phoneme sequence	Cross correlation
		Die aufgepflanzten Segel wurden größer und kamen näher: Hanake fühlte eine Bangigkeit	di aʊfʒəplantstən zegəl vʊrdəŋfrosz ʊnt kamən naər hanakə fultə aɪnə baŋɪkait	0.6267
		als müßte die Leinwand den Seeschaum reiben	als müstə di lainvant den zeʃaum raibən	0.6953
		Da trat ihre Lieblingsmagd die Singende Seemuschel herein und brachte einen zusammengerollten Brief	da trat irə liplɪŋsmakt di zɪŋfəndə zemʊʃəl herain ʊnt braχtə aɪnən tsuzamənʃərlən brif	0.6412

Table 4: Objective evaluation of German synthesizer.

Language: Dutch				
Target audio	Synthesized audio	Text sentence	Phoneme sequence	Cross correlation
		vervolgens de verhalen van den Noorweegschen bisschop Pontoppidan	vərvɔləŋs də vərhalə vən den norveχsən bisχɔp pɔntɔppidan	0.6101
		en onderzochten nu de zee in de rondte met de grootste nauwkeurigheid	ən əndərzoχtən ny də ze m də rəntə met də χrotstə naukørøχheit	0.5682
		Doch deze gebeurtenis kreeg bijzonder groote ruchtbaarheid en door de herkomst van het schip waarmede het plaats vond	dɔχ dezə χəbørtənis krex̩ bɪzɔndər χrotə ruydbarheit en dor də herkømst van ət sχip varmedə ət plats vɔnt	0.6429

Table 5: Objective evaluation of Dutch synthesizer.

Language: Bengali				
Target audio	Synthesized audio	Text sentence	Phoneme sequence	Cross correlation
		পশ্চালন কেলেঙ্কারির সময় তিনি গোড়ায় টেজারি অফিসারের দায়িত্বে ছিলেন।	poʃupalon keleŋkarir ʃ̠moe tini god̠dai t̠rejari ɔfisarer daiit̠te c̠hilen	0.6465
		নিজের পড়ার ঘর থেকে প্রায়ই বাচ্চা মেয়েটির কান্নার আওয়াজ পেত শুতি।	nijer p̠orar g̠h̠or t̠heke praj̠i bac̠h̠a meeṭ̠ir kannar aqaj̠peto sruti	0.6641
		মহিলাদের প্রতি অশালীন ইশ্পিত করার প্রতিবাদ করেছিলেন শ্যামল চোঙ্দোর।	mohilader proti ɔʃalin ingit k̠orar protibad korec̠hilen f̠æmol t̠ʃoŋdar	0.5828

Table 6: Objective evaluation of Bengali synthesizer.

Espeak Synthesized Audio	IPA2WAV:PTS Synthesized audio	Espeak MOS Result	PA2WAV:PTS MOS Result
Language: English LJSpeech			
		2.3	3.7
		1.7	3.3
		1.7	4.0
Language: English CMUArctic male			
		1.7	2.7
		2.0	3.3
		2.3	3.7
Language: English CMUArctic female			
		1.7	3.0
		2.0	4.0
		2.3	4.3
Language: German			
		1.6	3.7
		1.0	4.3
		1.0	3.7
Language: Dutch			
		1.3	3.7
		1.0	3.0
		1.6	3.6
Language: Bengali			
		1.3	3.0
		1.0	3.3
		1.6	4.0
Average MOS Score		1.6166	3.5722

Table 7: Subjective evaluation of PTS synthesizer compared to Espeak for different languages.

