

## Appendix A: Evaluation of TTS engines on our corpus (TSD paper number 1023)

This appendix contains the details on how each engine fared against each test case. In the tables below, each table corresponding to each test case, contain the problems related to mispronunciations we observed among the engines. Thus, for example, in Test Case 1, Mimic pronounces 14 words (from a sentence that contains 18 words) correctly, leading to an  $\phi$  score of 0.78. It mispronounced four words as shown in the rightmost three columns of the first two rows.

There was some variation demonstrated by the engines as they pronounced a word, or a set of words in the sentence; this variation is documented as follows. When the TTS engine did not pronounce a word, we denote that event by a “silence” in the table. When the TTS engine pronounced a word, but we could not understand the pronunciation, we document that event as “inaudible”. In cases where the TTS engine is unable to pronounce the word, it may render a phonetic realization instead. For example, “Sq.” may be rendered as the phonetic equivalent of the letters “s” and “q”, which is documented in the table cell as “es” “kyu”. A “long form” in a table cell implies that the engine renders a zip code of 60654 as “sixty thousand, six hundred and sixty four”.

(The tables begin on the next page.)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	14	0.78	&: ampersand	NW: en dɒbɔlyu	Apt.: apt
			Inc.: inc		
MaryTTS	12	0.67	Inc.: inc	NW: en dɒbɔlyu	Apt.: apt
			1A: 1-eh	MN: em en	60665: long form
FLITE	14	0.78	&: ampersand	Inc.: Inc	NW: en dɒbɔlyu
			Apt.: apt		
Voicery	11	0.61	Inc.: inc	NW: en dɒbɔlyu	St.: saint
			Apt.: apt	1A.: 1-amper	MN.: em en
			60665: long form		
Selvy	15	0.83	NW: en dɒbɔlyu	St.: saint	Apt.: apt
Acapela	18	1.00			
DeepVoice3	11	0.61	&: inaudible	Inc.: inaudible	NW.: new
			St.: saint	Apt.: apt	MN.: min
			60655: inaudible		
Polly	18	1.00			

Table 1: Test case 1 — American Communications & Engineering, Inc. is located at 123 NW. Main St., Apt. 1A, St. Paul, MN 60655 (18 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	12	0.67	Ctr.: si ti ar	Pkwy.: pi ke dɒbəlyu way	S.: es
			Ste.: suit	Ft.: ef ti	FL.: ef el
MaryTTS	11	0.61	Ctr. inaudible	Pkwy.: pi ke dɒbəlyu way	S.: es
			Ft.: ef ti	FL.: ef el	
FLITE	13	0.72	Ctr.: si ti ar	Pkwy.: pi ke dɒbəlyu way	Ste.: suit
			S.: es		
Voicery	12	0.67	Ctr.: si ti ar	S.: es	Ste.: see ti
			Ft.: ef ti	FL.: ef el	
Selvy	16	0.89	S.: es	Ft.: feet	
Acapela	18	1.00			
DeepVoice3	13	0.72	Ctr.: inaudible	Pkwy.: inaudible	Ste.: state
			FL.: silence		
Polly	18	1.00			

Table 2: Test case 2 — Natoma Professional Ctr. 555 Oakdale Pkwy., is located at 123 S. 2nd Pkwy., Ste. 700, Ft. Lauderdale, FL. (18 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	7	0.58	Ltd.: el ti di	Pl.: pi el	P.O.B.: pi o bi
			Sq.: es kyu	E.: e	
MaryTTS	5	0.42	1910: long form	Pl.: pi el	P.O.B.: pi o bi
			93425: long form	Sq.: es kyu	CA.: cah
FLITE	7	0.58	Ltd.: el ti di	Pl.: pi el	P.O.B.: pi o bi
			Sq.: es kyu	E.: e	
Voicery	7	0.58	E.: e	Pl.: pi el	P.O.B.: pi o bi
			93425: long form	CA.: cah	
Selvy	10	0.83	93425: long form	CA.: cah	
Acapela	11	0.92	E.: e		
DeepVoice3	6	0.50	E.: e	Pl.: pull	P.O.B.: silence
			93425: long form	Sq.: inaudible	CA.: cah
Polly	10	0.83	Pl.: pi el	CA.: cah	

Table 3: Test case 3 — Valor Telecom Ltd. 1910 E. Kimberly Pl. P.O.B 93425, Old Village Sq., CA. (28 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	14	0.82	Sec.: sec	Sen.: sen	Pres.: prez
			Sq.: es kyu		
MaryTTS	13	0.76	Sec.: sec	Sen.: sen	Pres.: prez
			FASB: fhas		
FLITE	14	0.82	Sec.: sec	Sen.: sen	Pres.: prez
Voicery	14	0.82	Sec.: sec	Pres.: prez	FASB: ef ay es bi
Selvy	17	1.00			
Acapela	17	1.00			
DeepVoice3	13	0.76	Sec.: sec	Sen.: sen	Pres.: prez
			FASB: fhas		
Polly	16	0.94	FASB: fahas bi		

Table 4: Test case 4 — Sec. of State Hillary Clinton and Sen. Lisa Murkowski spoke with Pres. Mahmaud Abbas to discuss FASB. (17 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	16	0.89	Ex-Gov: ex-gov	CEO: chee-o	
MaryTTS	16	0.89	Ex-Gov: ex-gov	RBOC: ar-bhak	
FLITE	17	0.94	CEO: chee-o		
Voicery	17	0.94	Ex-Gov: ex-gov		
Selvy	18	1.00			
Acapela	18	1.00			
DeepVoice3	14	0.78	Ex-Gov: ex-gov	Ex-HP: inaudible	CEO: inaudible
			RBOC: inaudible		
Polly	17	0.94	RBOC: ar-bhak		

Table 5: Test case 5 — Ex-Gov. Sarah Palin and Ex-HP CEO Carly Fiorina met with Israeli Ex-Prime Minister Ariel Sharon to discuss RBOC. (18 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations	
Mimic	13	0.72	Sec.: sec	Geithner: githner
			JPMorgan: jay pi em organ	\$4.5-million-a-year: point- a year
MaryTTS	13	0.72	Sec.: sec	\$4.5-million-a-year: four d and fifty cent
			JPMorgan: jay pi em organ	from Harvard: inaudible
FLITE	13	0.72	Sec.: sec	Geithner: geethner
			JPMorgan: jay pi em oh ar gee ay en	\$4.5-million-a-year: four p million a year
Voicery	15	0.83	Sec.: sec	COO: coo
			\$4.5-million-a-year: four dollar and five cent million a year	
Selvy	17	1.94	\$4.5-million-a-year: 4 point 5 dollars million a year	
Acapela	18	1.00		
DeepVoice3	11	0.61	Sec.: sec	F.: inaudible
			used to be the: inaudible	COO: coo
			\$4.5-million-a-year: four dollar and five cent million a year	MBA: emba
Polly	17	0.94	COO: coo	

Table 6: Test case 6 — Treasury Sec. Timothy F. Geithner used to be the COO at JPMorgan and earned \$4.5-million-a-year and earned an MBA from Harvard. (18 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	7	0.88	Rep.: rep		
MaryTTS	7	0.88	Rep.: rep		
FLITE	7	0.88	Rep.: rep		
Voicery	7	0.88	Chas.: chas		
Selvy	8	1.00			
Acapela	8	1.00			
DeepVoice3	5	0.63	Chas.: chas	Ph.D.: inaudible	PETA: pete
Polly	8	1.00			

Table 7: Test case 7 — Rep. Chas. Rangel Ph.D was censured by PETA. (8 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	20	0.80	Sr.: es ar	&.: ampersand	Co.: co
			Jr.: jay ar	Ms.: em es	
MaryTTS	23	0.92	Sr.: inaudible	Jr.: inaudible	
FLITE	20	0.80	Sr.: es ar	&.: ampersand	Co.: co
			Jr.: jay ar	Ms.: em es	
Voicery	24	0.96	Co.: co		
Selvy	25	1.00			
Acapela	25	1.00			
DeepVoice3	23	0.92	Sr.: sir	&.: silence	
Polly	25	1.00			

Table 8: Test case 8 — Mr. John Smith Sr. and Mrs. Jane Smith worked at Levi Strauss & Co. with their son, John Smith Jr., and daughter Ms. Judy Smith. (25 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	10	0.91	NYSE: neighs		
MaryTTS	9	0.82	Gen.: gen	NYSE: neighs	
FLITE	10	0.91	NYSE: neighs		
Voicery	10	0.91	Gen.: gen		
Selvy	11	1.00			
Acapela	11	1.00			
DeepVoice3	10	0.91	NYSE: neighs		
Polly	11	1.00			

Table 9: Test case 9 — Gen. Douglas MacArthur was tired of receiving SPAM from the NYSE. (11 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	10	1.00			
MaryTTS	10	1.00			
FLITE	10	1.00			
Voicery	10	1.00			
Selvy	10	1.00			
Acapela	10	1.00			
DeepVoice3	10	1.00			
Polly	10	1.00			

Table 10: Test case 10 — They were too close to the door to close it. (10 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	5	0.83	dove (second instance): duv		
MaryTTS	5	0.83	dove (second instance): duv		
FLITE	5	0.83	dove (second instance): duv		
Voicery	4	1.67	dove (first instance): dev	dove (second instance): dev	
Selvy	6	1.00			
Acapela	6	1.00			
DeepVoice3	5	0.83	dove (first instance): dev		
Polly	6	1.00			

Table 11: Test case 11 — The dove dove into the water. (6 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	7	0.88	lead: led		
MaryTTS	7	0.88	lead: led		
FLITE	7	0.88	lead: led		
Voicery	6	0.75	lead: led	lead: leed	
Selvy	7	0.88	lead: leed		
Acapela	7	0.88	lead: leed		
DeepVoice3	7	0.88	lead: leed		
Polly	7	0.88	lead: leed		

Table 12: Test case 12 — The team lead had lead us to victory. (8 words)



TTS Engine	Correct Words	$\phi$	Problems and mispronunciations
Mimic	15	0.94	read: red
MaryTTS	15	0.94	read: red
FLITE	15	0.94	read: red
Voicery	16	1.00	
Selvy	16	1.00	
Acapela	16	1.00	
DeepVoice3	15	0.94	read: reed
Polly	15	0.94	read: red

Table 13: Test case 13 — After I read a book I add it to my list of books that I’ve read. (16 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations
Mimic	6	0.86	produce: prədus
MaryTTS	6	0.86	produce: prədus
FLITE	6	0.86	produce: prədus
Voicery	7	1.00	
Selvy	6	1.00	produce: prədus
Acapela	6	1.00	produce: prədus
DeepVoice3	6	0.86	produce: prədus
Polly	7	1.00	

Table 14: Test case 14 — The farm was used to produce produce. (7 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	6	0.86	use: juz		
MaryTTS	7	1.00			
FLITE	7	1.00			
Voicery	7	1.00			
Selvy	7	1.00			
Acapela	7	1.00			
DeepVoice3	7	1.00			
Polly	7	1.00			

Table 15: Test case 15 — People who use are of no use. (7 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	8	0.80	Prof.: proff	Robt.: robt	
MaryTTS	8	0.80	Prof.: proff	Robt.: robt	
FLITE	8	0.80	Prof.: proff	Robt.: robt	
Voicery	8	0.80	Prof.: proff	Robt.: robt	
Selvy	10	1.00			
Acapela	10	1.00			
DeepVoice3	6	0.60	Prof.: proff	Robt.: robt	B.: null
			Reich: rich		
Polly	9	0.90	bona fide: bona fidee		

Table 16: Test case 16 — Prof. Robt. B. Reich is a bona fide rocket scientist. (10 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	12	0.92	chutzpah: chut-ze-pah		
MaryTTS	12	0.92	chutzpah: chut-ze-pah		
FLITE	12	0.92	chutzpah: chut-ze-pah		
Voicery	12	0.92	chutzpah: chut-ze-pah		
Selvy	13	1.00			
Acapela	13	1.00			
DeepVoice3	10	0.77	Einstein: eensteen	PhD: silence	
Polly	13	1.00			

Table 17: Test case 17 — Dr. Albert Einstein, PhD had a lot of chutzpah turning down the presidency. (13 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	13	0.81	III: ay ay ay	Capt.: cap't	Wm: døbəlyu er
MaryTTS	12	0.75	A.: ah	III: inaudible	Capt.: cap't
			W.: inaudible		
FLITE	13	0.81	III: ay ay ay	Capt.: cap't	Wm: døbəlyu er
Voicery	12	0.75	Jas.: jas	A.: ah	III: three
			coup d'etat: coo det at		
Selvy	15	0.94	III: three		
Acapela	15	0.94	III: three		
DeepVoice3	11	0.69	Jas.: jas	III: ay ay ay	Wm.: inaudible
			O.: inaudible	coup d'etat: cup detet	
Polly	16	1.00			

Table 18: Test case 18 — Jas. A. Barone III said bon voyage to Capt. Wm. O. Barnett before the coup d'etat. (16 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	11	0.73	Mon.: moan	Sept.: sep't	
MaryTTS	11	0.73	Mon.: moan	Sept.: sep't	25: twenty-five
FLITE	11	0.73	Mon.: moan	Sept.: sep't	
Voicery	10	0.67	AM: ay		
Selvy	13	0.87	Mon.: mun		
Acapela	13	0.87			
DeepVoice3	10	0.60	Mon.: moan	Sept.: sep't	25: twenty-five
			AM: am		
Polly	15	1.00			

Table 19: Test case 19 — I was born Mon., Sept. 25, 1989 at 12:30 AM. (10 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations		
Mimic	11	0.73	Lt.: el ti	Cmd.: cee em dee	Wed.: wed
			Mar.: mar		
MaryTTS	11	0.73	Lt.: inaudible	Cmd.: inaudible	Wed.: wed
			Mar.: mar		
FLITE	11	0.73	Lt.: el ti	Cmd.: cee em dee	Wed.: wed
			Mar.: mar		
Voicery	10	0.67	Lt.: el ti	Cmd.: cee em dee	Jas.: jas
			Wed.: wed	3rd.: three	
Selvy	13	0.87	Wed.: wed	Mar.: mar	
Acapela	13	0.87	Wed.: wed	Mar.: mar	
DeepVoice3	10	0.67	Cmd.: inaudible	Jas.: jas	W.: inaudible
			Wed.: wed	Mar.: mar	
Polly	15	1.00			

Table 20: Test case 20 — Lt. Cmd. Jas. W. Marks was born Wed. the 3rd. Of Mar. at 2:30 p.m. (15 words)

TTS Engine	Correct Words	$\phi$	Problems and mispronunciations
Mimic	14	1.00	
MaryTTS	14	1.00	
FLITE	14	1.00	
Voicery	14	1.00	
Selvy	14	1.00	
Acapela	14	1.00	
DeepVoice3	13	0.93	Eau Claire: ew
Polly	14	1.00	

Table 21: Test case 21 — I live in La Crosse county, Wisconsin. This is close to Eau Claire and Prarie du Chien. (14 words)