Appendix A: Speech Synthesis Demonstration CD

A Historical Review Collected by Dennis Klatt.

Dennis Klatt (1987). Review of text-to-speech conversion for English, Journal of the Acoustical Society of America, 82 (3), pp. 737-793.

Part A: Development of speech synthesizers

01.	VODER, Homer Dudley, 1939. "Good evening radio audioence" ... "Good afternoon radio audioence".	0.18
02.	The Haskins Laboratories Pattern Playback by Franklin Cooper, 1951.	0.11
03.	PAT (Parametric Artificial Talker) parallel formant synthesizer designed by Walter Lawrence, 1953.	0.06
04.	OVE (Orator Verbis Electris) cascade formant synthesizer, Gunnar Fant, 1953.	0.04
05.	Copying natural sentence using PAT, 1962.	0.05
06.	Copying same sentence using the second generation of OVE, 1962.	0.05
07.	Comparison of synthesis and a natural sentence using OVE II, John Holmes, 1961.	0.09
08.	Comparison of synthesis and a natural sentence, John Holmes using his parallel formant synthesizer, 1973.	0.07
09.	Attempt to scale the DECtalk male (Perfect Paul) voice to make it sound female.	0.12
10.	Comparison of synthesis and a natural sentence, female voice, Dennis Klatt, 1986.	0.10
11.	The DAVO articulatory synthesizer, George Rosen, MIT, 1958.	0.28
12.	Sentences produced by an articulatory model, James Flanagan and Ishizaka, 1976.	0.04
13.	Linear-prediction analysis and resynthesis of speech at low bit-rate, Texas Instruments Speak'n'Spell toy, Richard Wiggins, 1980.	0.12
14.	Comparison of synthesis and a natural recording, automatic analysis-resynthesis using multipulse linear prediction, Bishnu Atal, 1982.	0.09

Part B: Segmental synthesis by rule

15.	Creation of a sentence from rules in the head of Pierre Delattre, using the Haskins Pattern Playback, 1959.	0.06
16.	The first computer-based phonemic synthesis-by-rule program, John Kelly and Louis Gerstman, 1961.	0.12
17.	Elegant rule program for British English by John Holmes, Ignatius Mattingly, and John Shearme, 1964.	0.11
18.	Formant synthesis using diphone concatenation, Red Dixon and David Maxey, 1968.	0.14
19.	Rules to control a low-dimensionality articulatory model, Cecil Coker, 1968.	0.15

Part C: Synthesis by rule of sentences and sentence prosody

20.	First prosodic synthesis by rule, Ignatius Mattingly, 1968.	0.22
21.	Sentence level phonology incorporated in rules by Dennis Klatt.	0.12
22.	Concatenation of linear-prediction diphones, Joe Olive, 1977.	0.13
23.	Concatenation of linear-prediction demisyllables, Catherine Brownman, 1980.	0.18

Part D: Fully automatic text-to-speech conversion

24.	The first full TTS system, Noriko Umeda et al., 1968.	0.19
25.	The first Bell Laboratories TTS system, Cecil Coker, Noriko Umeda, and Catherine Browman, 1973.	0.16
26.	The Haskins Laboratories TTS system, 1973.	0.19

27.	The Kurzweil reading machine for the blind, Raymond Kurzweil, 1976.	0.15
28.	The inexpensive Votrax Type'n'Talk system, Richard Gagnon, 1978.	0.07
29.	The Echo low-cost diphone concatenation system, 1982.	0.14
30.	The M.I.T. MITalk system by Jonathan Allen, Sheri Hunnicutt, and Dennis Klatt.	0.17
31.	The multi-language Infovox system by Rolf Calrson, Bjorn Granström, And Sheri Hunnicutt, 1982.	0.15
32.	The Speech Plus Inc. Prose-2000 commercial system, 1982.	0.11
33.	The Klattalk system, Dennis Klatt, 1983.	0.24
34.	AT&T Bell Laboratories TTS system, 1985.	0.25
35.	Several of the DECtalk voices: Perfect Paul, Beautiful Betty, Huge Harry, Kit the Kid, and Whispering Wendy.	0.33
36.	DECtalk speaking at about 300 words/minute.	0.12

DECtalk PC (Digital Equipment Corporation)

37.	An improved version of DECtalk. "DecTalk PC is a new and improved version of the well-known DecTalk speech synthesizer. It represents the state of the art in text-to-speech synthesis. This demonstration will provide some examples of DecTalk's capabilities." "DECtalk has nine standard voices and can speak at rates of 75 to 600 words per minute." "The DECtalk text-to-speech synthesizer converts computer text to intelligible speech."	0.29

Eurovocs TTS system (Technologie & Revalidatie)

38.	American English. "PCSes are coming. Today, a new form of wireless services is generating a lot of excitement. While cellular digital packet data, packet radio, and eventually digital cellular will play the dominant roles in long-distance wireless communication, a new set of services called Personal communication Services could change the face of short-distance wireless."	0.27
39.	German. "Wirtschafts Wunder Wurst. Sie ist 15 Zentimeter lang, hat etwa drei Zentimeter Durchmesser, knapp 500 Kalorien und wiegt ungefähr 60 Gramm. Als Delikatesse kann man sie nicht wirklich bezeichnen. Aber was für den einen der Hamburger ist, ist für den anderen die Currywurst. Manche können nur die Nase rümpfen über die Wurst, die vor 45 Jahren in Berlin das Licht der Welt erblickte"	0.30
40.	French. "Eurovocs est capable de transformer un texte écrit en parole. La transposition en parole est caractérisée par certains paramètres. Eurovocs vous permets de modifier ces paramètres. A part cela, Eurovocs vous offres d'autres facilités. La mise en service des fonctions spéciales se réalise par l'envoi des commandes au synthétiseur. Ceci se passe toujours selon le même mode"	0.26

An Improved Version of Eurovocs, Released in 1995

41.	American English. "Hi! Welcome to our demonstration of the American English TTS system. This program converts any written text into a phonetic representation. From this representation, a spoken version is then synthesized. Thus, the artificial voice which you have just been listening to. Thank you for your attention. Bye."	0.24
42.	German. "Guten Tag, und willkommen bei der Demonstration unseres deutschen Text nach Sprach Systems. Sie hören eine syntheetische Stimme. Dieses Programm setzt zunächst beliebige Texte in eine phonetische Transkription um. Anschliessend wird mittels Diphon-Synthese eine gesprochene Fassung des Textes erstellt. Wir danken Ihnen für Ihr Interesse. Auf Wiedersehen."	0.23

43. French. " Bonjour! Bienvenue ànotre démonstration du système français de synthèse de la parole. Ce programme convertit tout texte français écrit en sa transcription phonétique. Le synthétiseur transforme ensuite cette transcription phonétique en parole. Ainsi a été produite la voix artificielle que vous entendez actuellement. Merci beaucoup de votre attention. Au revoir!"   0.23

44. Dutch. "Welkom bij deze demonstratie van ons nederlandstalig tekst-naar-spraak-systeem. U hoort een synthetische stem. Dit systeem zet om het even welke tekst om in een fonetische transcriptie. Vervolgens wordt op grond van deze transcriptie een gesproken boodschap gesynthetiseerd. Bedankt voor uw aandacht. Tot ziens."   0.24

45. Spanish. "Hola, bienvenidos a nuestra demostración para el idioma español. Este programa convierte cualquier texto escrito en una transcripción fonética. A partir de esta transcripción se sintetiza una versión hablada. Así, por ejemplo, la voz artificial que usted acaba de escuchar. Muchas gracias por su atención. Hasta pronto."   0.23

Telia Promotor Infovox TTS System (230 v1.0)

46. American and British English male voices.   0.28
47. Finnish male voice.   0.12
48. German male voice.   0.13
49. Italian male voice.   0.17
50. Spanish male voice.   0.17
51. French male voice.   0.17
52. Swedish, Norwegian, Icelandic, and Danish male voices.   0.59
53. Dutch male voice (230 v1.1).   0.18

Telia Promotor Infovox TTS System (330 v1.0)

54. British English male, German female, and Dutch male voices.   1.21

British Telecom BT-Laboratories Laureate TTS System

55. English male (Southern British accent). "The 5:30 train for Liverpool St. is running 7 minutes early."
English female (Southern British accent). "Peter Piper picked a peck of pickled herring" ... "There are 3 people listed under Johnston, Mr..."
English male (Northern accent). "This is synthetic speech based on the voice of Peter Cochrane".
Dialogue between a male and female speaker. "My wife has just gone to the West Indies..."   0.28

Bellcore (Bell Communications Research) ORATOR TTS System

56. General text and names. "Cat lives 9 lives. Close the book after a close look. Be content of the content of the course." ... "Rosemary Ellington, Christoffer Robert Hallingsworth, Kathleen McIntyre, John Pennington, Nancy Weaver, Fred Flintstone, Tony Baldaccini, Thadeus Bialobrzeski, Said Habib Jeshaia Schnitzer, Tsuyoshi Utsunomiya, Theodoros Xanthopoulos."   0.30

57. An improved ORATOR 2. "Hello. My name is ORATOR 2. I am a speech synthesizer. I convert English text to speech. Address e-mail to: orator@bellcore.com"   0.15

AT&T Bell Laboratories (Lucent Technologies) TTS System

58. English greeting with male, female, child, big man, and ridiculous voices.
"Welcome to the Bell Labs text-to-speech system."
English abbreviations and numbers: "Lumber will cost $3.95 for 7 ft. on Sat."
"That fossil from NM is 165,256,011 yr. old" "Dr. Smith lives on Oak Dr., but St.
John lives on 71st St."   0.36

59. German greeting.   0.07

60. French greeting.   0.10

61. Italian greeting.   0.18

62. Spanish greeting.   0.15

63. Chinese greeting.   0.12

64. English song: 'Bicycle made for two'   0.32

ELAN Informatique ProVerbe Speech Engine

65. American English male and female voices. British English male voice.   0.10

66. French male and female voice.   0.08

67. German male voice.   0.06

68. Spanish male voice.   0.05

SoftVoice Inc. SVTTS System

69. Male voice. "We have all the flavour, as long as you want chocolate, vanilla, or
strawberry."
Breathy female voice. "I am here to help you. Please enter your user number."
Child voice. "Twinkle twinkle little star. How I wonder what you are."
Colossus voice. "This is the voice of world control. Obey me and live, or disobey
and die"
Various "I woke this morning with fairly good idea who I was, but I quess I was
mistaken."   0.30

MBROLA Diphone Based Non-TTS System

70. British English male and American English female voices.   0.18

71. German male and female voices.   0.13

72. French male and female voices.   0.14

73. Dutch, Portuguese, Spanish, and Romanian male voices.   0.23

ETI Eloquence

74. Male voice "Hello, my name is Wade. I am an adult male."
Male voice "Hello, my name is Glen. I am Wade's friend, but I'm much breathier."
Female voice "Hi, my name is Lou. I am an adult female, in case you didn't
know."
Male and female voices: "Today is a spectacular day."   0.19

AcuVoice Concatenative Speech Synthesizer

75. English male voice and names.   0.55

Microsoft Whistler Trainable TTS System

76. English male voice.   0.13

Festival Diphone Synthesis System (CSTR, University of Edinburgh)

77.	English male voice.	0.36
78.	SSML (Speech Synthesis Markup Language) demonstration.	0.25

Lyricos Singing Voice System

79.	Singing and choir.	0.20

HMM Based Speech Synthesis by Robert Donovan

80.	Male voice: "My name's Rob Donovan, and I'm from Cambridge University" ... "This is an example of the best speech produced by the system to date. It was trained on text read from a novel, namely the Hitch Hiker's Guide to the Galaxy. The Modified Rhyme Test error for this speech is only five point zero percent. I must apologise for speaking on a monotone, but you see, I have no brain." Female voice: "I don't normally talk in a monotone you know! Rob made me do it!" Synthetic male voice from natural female voice: "In the beginning was the word, and the word was with God, and the word was God."	0.39

L&H (Learnout & Hauspie) BeST TTS System

81.	English male and female, Dutch male, Japanese Female, and Chinese Female voices.	0.19

Listen2

82.	English male with statement, question, and pretty voices. English female voice.	0.15
83.	German, French, Italian, and Spanish male voices.	0.16

ASEL ModelTalker (University of Delaware)

84.	Male voice with normal (long), happy, sad, frustration, and assertive. "The ASEL ModelTalker TTS system converts plain English text to speech. It uses a text to phoneme system which includes capabilities for parsing ToBI-like descriptions of the intonation, and a diphone-based phoneme to sound engine."	0.42
85.	Female and child voices with the same emotional states. (experimental diphone inventory). "When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow."	1.12

JSRU Parallel Formant Synthesizer

86.	Male voice. "John would give an utterance." x 3	0.08

Apple PlainTalk

87.	MacinTalk Pro. Male (Bruce) and female (Victoria) voices. MacinTalk 3. Male, female, and child voices. Singing good news, bad news, organ, and cello. MacinTalk 2. Male, female, and child voices.	1.11

Panasonic CyberTalk

88.	Male and Female voices. "The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down a very deep well."	0.25

SPRUCE High-level System

89. Weather Forecast.   0.09

HADIFIX German Speech Synthesizer (University of Bonn)

90. German female voice.   0.13

SVOX German Speech Synthsizer

91. German male voice.   0.47

CHATR Speech Synthesis System

92. English male and female voices.   0.23

SYNTE2 Finnish Speech Synthesizer

93. Finnish male voice: "Minä olen Tampereella kehitetty suomea puhuva
     puhesyntetisaattori. Minun nimeni on SYNTE2. Vaikka minut alunperin
     kehitettiinkin pääasiallisesti vammaisten kommunikaatioapuvälineeksi, voidaan
     minua käyttää moniin muihinkin tarkoituksiin. Näitä voisivat olla esimerkiksi
     automaattiset kuulutusjärjestelmät lento- ja rautatieasemilla,
     prosessinvalvontalaitteistot teollisuudessa, tietokoneiden puhetulostus ja niin
     edelleen."   0.44

Timehouse Mikropuhe v4.11

94. Finnish male speech.   0.40
95. Singing voice: "Finlandia"   0.44

Sanosse

96. Normal text and numbers 1 ... 10.   0.25
97. Telephony application version by Sonera.   1.36

SYNTE3 Finnish Speech Synthesizer

98. Finnish male voice: "SYNTE3. 1, 2, 3, 4, 5, 123, 234, 4321."   0.19