



Integrating TTS and Recorded Speech

▶▶ *The best of both worlds*

Marc Fabiani







© Nuance Communications, Inc. 2007



NUANCE

© Nuance Communications, Inc. 2007

The State of the Art

- ▶ You've come a long way, baby: 
- ▶ Equivalency with recordings is possible
 - Recorded Speech 
 - TTS 
- ▶ Accuracy can exceed human ability with perfect consistency
 - ▶ Bexar, Texas 
 - ▶ Bagehot 
 - ▶ Natchitoches, Louisiana 
- ▶ TTS Voices are available in many languages
- ▶ Easy to Use (text in -> speech out)
- ▶ Cost effective

So why isn't TTS used everywhere?

- ▶ Variety
 - TTS offers at best a half dozen voices per language
 - Custom voices are expensive and time consuming
 - With recordings the choice is enormous
- ▶ Expressivity
 - TTS can recite but it can't act
- ▶ TTS will address these issues eventually, but until it does recordings will continue to be used

What's wrong with using recordings?



The Daily Record



United We Stand

What else is wrong with using recordings?

- ▶ Breaks the simplicity of Text in -> Speech out
- ▶ Complexifies application development
- ▶ Exposes platform and/or OS based inconsistencies

Why is this a big deal?

- Software as a Service is the future of TTS
- Some characteristics of Software as a Service:
 - Remotely Hosted, Network Accessed, Ubiquitous
 - Thin Client
 - Simple API
 - Transaction and/or subscription based
- This is not a new idea



Remotely Hosted – Network Accessed – Ubiquitous

THE LITERARY DIGEST MARCH 4, 1922

WITHIN THE REACH OF MILLIONS



The most valuable things on earth are the invisible things. Gifts of Wisdom—Dignity—good, courage and calm—in the life, good resolution and helpfulness, Gifts of human nature—Love, Sympathy, kindness and friendly speech.

There, of material things, some of the most useful are the commonest and cheapest. These we take almost for granted. There is no use to make their actual worth.

It is a great failure to the value of the telephone that within a few short generations it has come to be ranked among these common things. To date we are a habit of millions of people. It speaks and acts and simplifies living. It extends the range of your very presence. It allows you to give advice, assistance—a real advantage to those of need.

Daily it saves untold expense and worry, multiplies working power, saves attention, combines together the human fabric. Helps the individual man and woman to strength over the combination of a vast world.

You cannot make half the work of a world and answer a thing as the telephone. You can only know that the value must be infinite.

THE SATURDAY EVENING POST



Less than 5% is in your home!

men and women throughout the country. When you lift your receiver, the apparatus goes into action. Wire, cable, switchboards—thousands of them, big and little—all must be ready whenever you call.

And all this equipment must be made with such precision that your Bell telephone will work in perfect coordination with the millions of others in the country. What other product do you use but to meet such a test?

Suchler's apparatus of high quality is absolutely essential. To make sure of this, the Bell System has for 56 years entrusted its manufacturing to Western Electric.

Back of the Bell Telephone in your home is a vast network of equipment that you never see. Back of that are the telephone company, Western Electric and their thousands of trained




1. Look into one of the great Western Electric cable shops. In a year over six million miles of wire are put into these familiar cables.

2. These machines turn out rubber covered telephone wire. Each year the Company manufactures 300,000 miles of this important product.




3. Expert workmen install an intricate maze of wiring in these telephone switchboards at one of Western Electric's three great factories.

4. Five skilled workers are making delicate adjustments on that maze of intricacy, a dial type control effect.



Western Electric

BELL SYSTEM SERVICE IS BASED ON WESTERN ELECTRIC QUALITY

Simple API

"And I Can Make It Talk"

You're right about that, Sonny. At eight or eighty it's easy for any one to use the telephone. And year after year the service gets better and better.

This country leads the world in telephone service because it leads in telephone research. Thousands of scientists, engineers and assistants are at work in the Bell Telephone Laboratories to make the service faster, clearer and more economical. No part of the Bell System is more important to you as a telephone user.

BELL TELEPHONE SYSTEM

Your good telephone service is made possible by the constant research of Bell Telephone Laboratories

Calling Grandma looks so simple!

AND it is simple, whether you're making a call or receiving one. But it takes just about the most complex mechanism imaginable to make your calls possible.

Whenever you pick up the telephone, millions of Western Electric-made parts in the Bell System network are ready to go to work — to carry your voice clearly across the street or across the nation, as you may choose. You can see it is essential that all these parts work perfectly together. They do — and there's a good reason why they do:

• For 67 years Western Electric has been the manufacturing back of the Bell System. This has enabled one people who make telephone equipment to work most closely with Bell scientists who design it and Bell telephone people who operate it. Result: the best telephone service on earth — getting better and more valuable to you all the time.

Western Electric

A UNIT OF THE BELL SYSTEM SINCE 1882

THE, DECEMBER 14, 1953

Transaction and/or Subscription Based



**Pennies
For
Wings**

The voice of a friend. Reassuring words from father, mother, son or daughter. A hurried call for aid in the night. You cannot set a price on such things as these.

Yet this is true—telephone service is cheap in this country. No other people get so much service, and such good and courteous service, at such low cost.

BELL TELEPHONE SYSTEM

The Bell System cordially invites you to visit its exhibits at the New York World's Fair and the Golden Gate International Exposition, San Francisco.

JULY, 1939

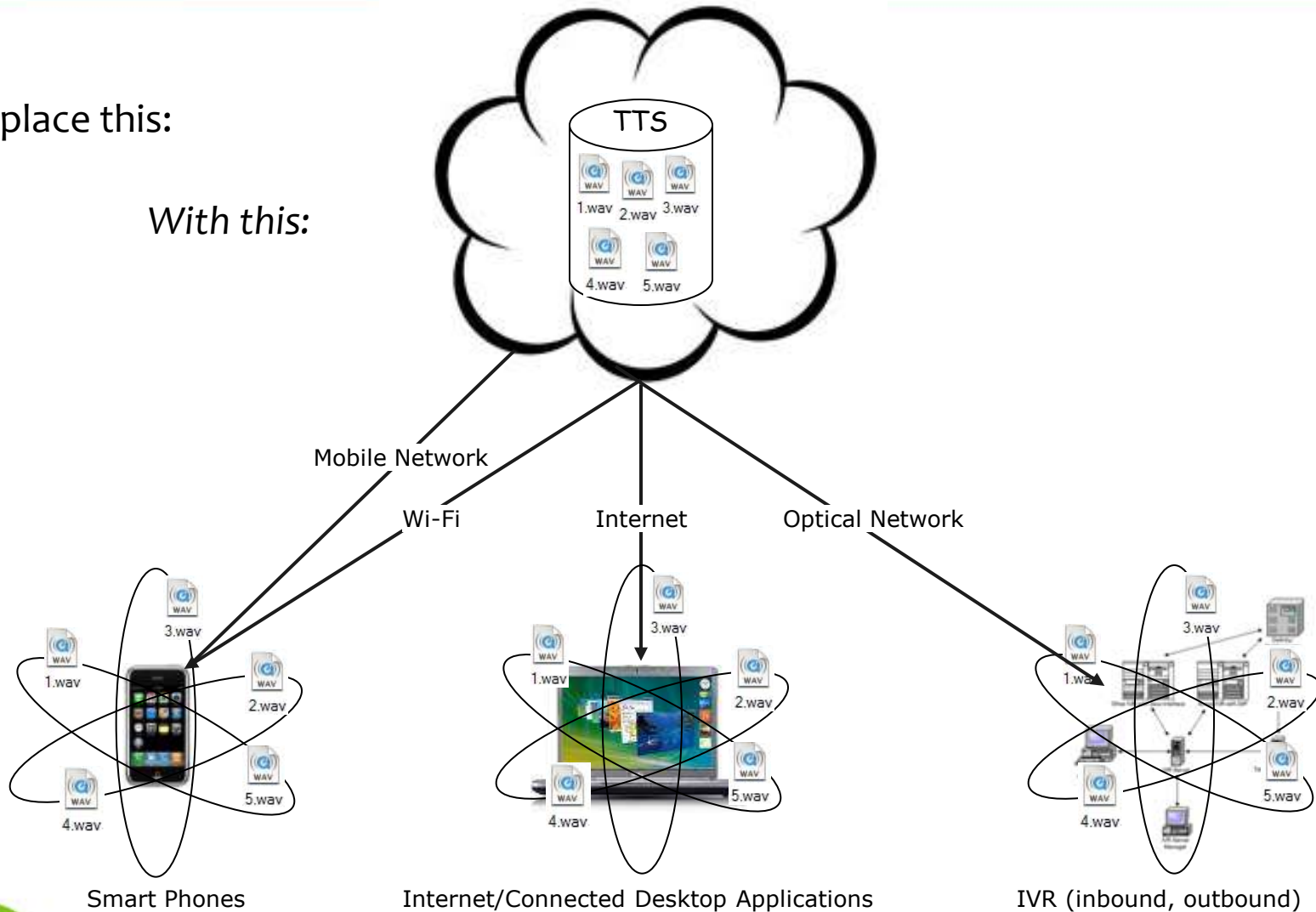
Photo courtesy of the Library of Congress

5




Solution

Replace this:

With this:



How?

- ▶ Custom Voices
- ▶ Partial Custom Voices
 - Provide pre-defined contexts for concatenating recordings
 - Dates
 - Times
 - Currency
 - Natural Number
 - Etc.
 - User simply drops in the recordings and the TTS engine does the rest
 - Text in -> Speech out 
 - Allow for the integration of *any* recordings  

▶▶ Live Demo