

디지털시대의 음성언어자료 이해 및 활용

Leveraging the Potential of Spoken Data in the AI-Powered Digital Age

Miran Kim (mirankim@gnu.ac.kr)

Dept. of English Education, Gyeongsang National University



• 주최: 연세대학교 언어정보연구원

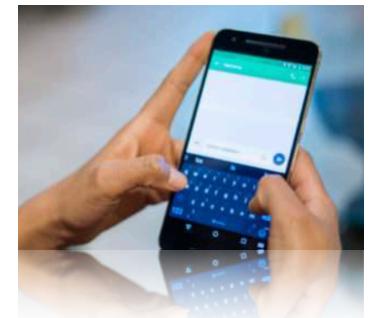
139회 학술발표회 2023. 1. 26 (목) 14:00 ~

Written language in the digital age



- The use of written language has increased significantly:

e.g., instant messaging, email, text messages, and social media posts.



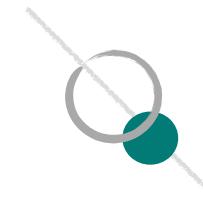
- More efficient and **effective communication** across long distances and time zones.
- New forms of written language: e.g., emoticons, acronyms and slang

DM

RT

별다줄

Spoken vs. Written language



Differences?

Transient	Permanent
Heavily rely on nonverbal cues	Words themselves
More informal / colloquial	More formal / polished
Cannot be easily edited / corrected	Edited / reviewed
In real-time communication	Delayed or asynchronous communication

e.g., Sumerian in ancient Mesopotamia around the 4th millennium BCE

Recording sound

- In the 4th century BCE, the Greek philosopher Aristotle wrote about **echoes** in his work "On the Heavens".



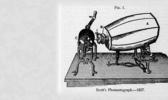
1

1859



Frenchman Édouard-Léon Scott de Martinville invents the first device capable of recording sound. Known as a **phonautograph**, the machine captures sound by tracing the vibrations of a bristle onto a sheet of soot-covered paper.

These recordings, called **phonautograms**, were never intended to be played back – Martinville's purpose in making the device was to simply see what sound waves looked like.



- The earliest known recording was made by Scott on April 9, 1860, it was a **recording of himself singing "Au clair de la lune" ("Under the Moon")**



2



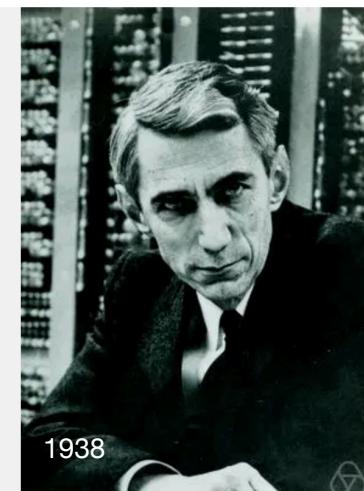
Photo of Edison with cylinder phonograph in 1878

Edison Invents the Phonograph

Thomas Edison created many inventions, but his favorite was the phonograph. While working on improvements to the telegraph and the telephone, Edison figured out a way to record sound on tinfoil-coated cylinders. In 1877, he created a machine with two needles: one for recording and one for playback. When Edison spoke into the mouthpiece, the sound vibrations of his voice would be indented onto the cylinder by the recording needle. What do you think were the first words that Edison spoke into the phonograph?



3

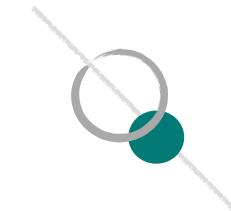


1938

Alec Reeves, PCM and the Birth of Digital Communication – historictech

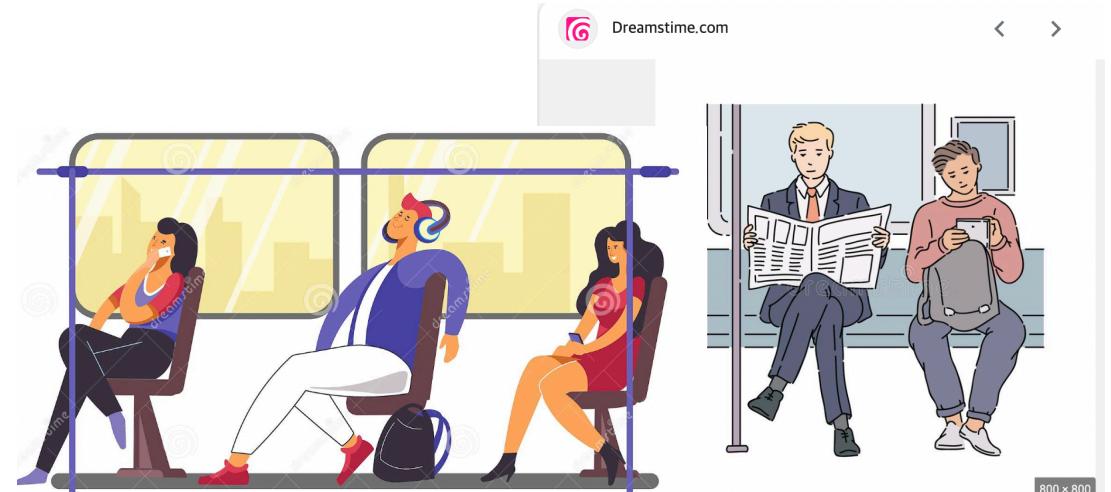


Spoken language in the digital age

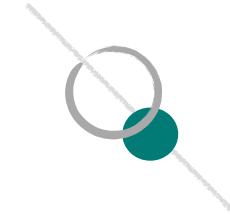


- Spoken language become increasingly important with the rise of **voice enabled technology**:

e.g, interaction with devices,
importance of audio contents
real-time interactions

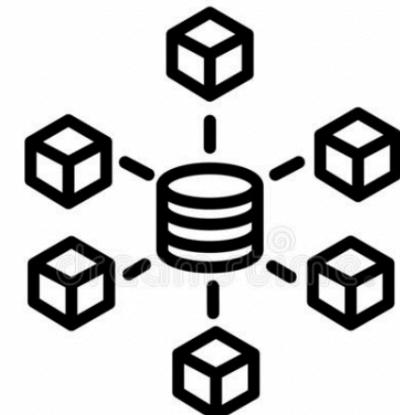


Two Men in Subway Reading Newspaper and Using Tablet
Stock Vector - Illustration of device, city: 147073734
images from dreamtime.com



Language & Information

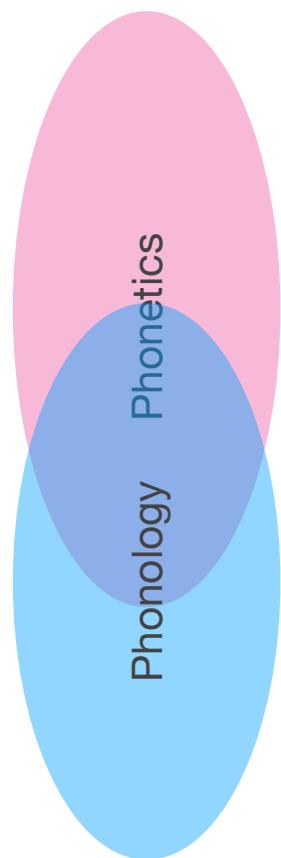
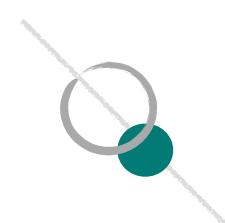
(written, spoken)



big data

Spoken language analysis

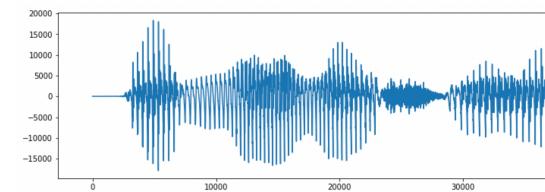
Phonetic vs. Phonological analysis



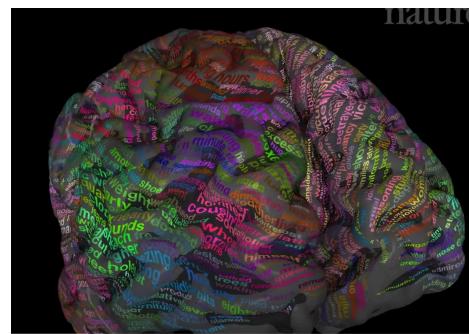
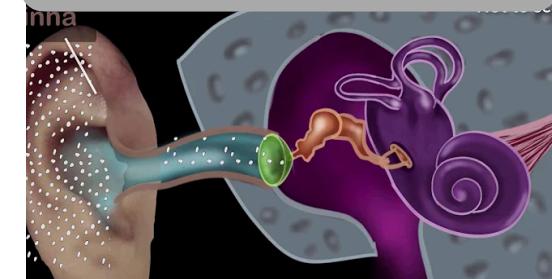
Articulatory Phonetics



Acoustic Phonetics



Auditory Phonetics



SEOUL Corpus: The Korean Corpus of Spontaneous Speech

- 한국어 자연발화 음성 코퍼스 (Yun, et al. 2015)

- Interview speech recordings of 40 Seoul dialect speakersK (20F + 20M; 10s, 20s, 30s, 40s)
- Total 240 sound files + 240 label files: e.g., 1 hr per speaker (6 speech files)

The screenshot shows the OpenSLR website interface. At the top, there's a navigation bar with 'OpenSLR' and 'Home Resources'. Below it, a section for 'SEOUL CORPUS' is displayed, including its identifier (SLR113), summary (The Korean Corpus of Spontaneous Speech), category (Speech), license (Attribution-NonCommercial 2.0 Generic), and download links for various file types (readme.tgz, label.tgz, sound.tgz). A detailed description of the resource follows, mentioning it was funded by the National Research Foundation of Korea and includes a BibTeX entry.

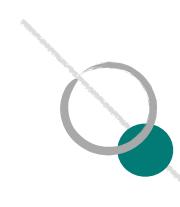
Identifier: SLR113
Summary: The Korean Corpus of Spontaneous Speech (aka, Seoul Corpus), created from the NRF(Korea)-funded project.
Category: Speech
License: Attribution-NonCommercial 2.0 Generic (CC BY-NC 2.0)
Downloads (use a mirror closer to you):
readme.tgz [27M] (readme, manual, search script, sample screenshot, journal paper) Mirrors: [\[US\]](#) [\[EU\]](#) [\[CN\]](#)
label.tgz [57M] (Praat TextGrid label files) Mirrors: [\[US\]](#) [\[EU\]](#) [\[CN\]](#)
sound.tgz [2.5G] (flac sound files) Mirrors: [\[US\]](#) [\[EU\]](#) [\[CN\]](#)

About this resource:
This data is phonetically labeled Korean Corpus of Spontaneous Speech (aka, Seoul Corpus). This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2012S1A5A2A03034027). You can cite the data using the following BibTeX entry:

```
@article{yunetal2015,  
    title={The Korean Corpus of Spontaneous Speech},  
    author={Yun, W., Yoon, K., Park, S., Lee, J., Cho, S., Kang, D., Byun, K., Hahn, H., Kim, J.J.},  
    journal={Journal of the Korean Society of Speech Sciences},  
    volume={7(2)},  
    pages={103 - 109},  
    year={2015},  
    publisher={}}}
```

- Recording period: 2012~2013
- 230,000 word tokens (51,000 types): 1.1M phoneme tokens
- Labeling (7 tiers): including phoneme, pronunciation form, syllables, words, transcription, etc.
- Used Forced alignment process
- Free distribution for research

Information transfer in the digital age



Input



Input



Input



Output



Output



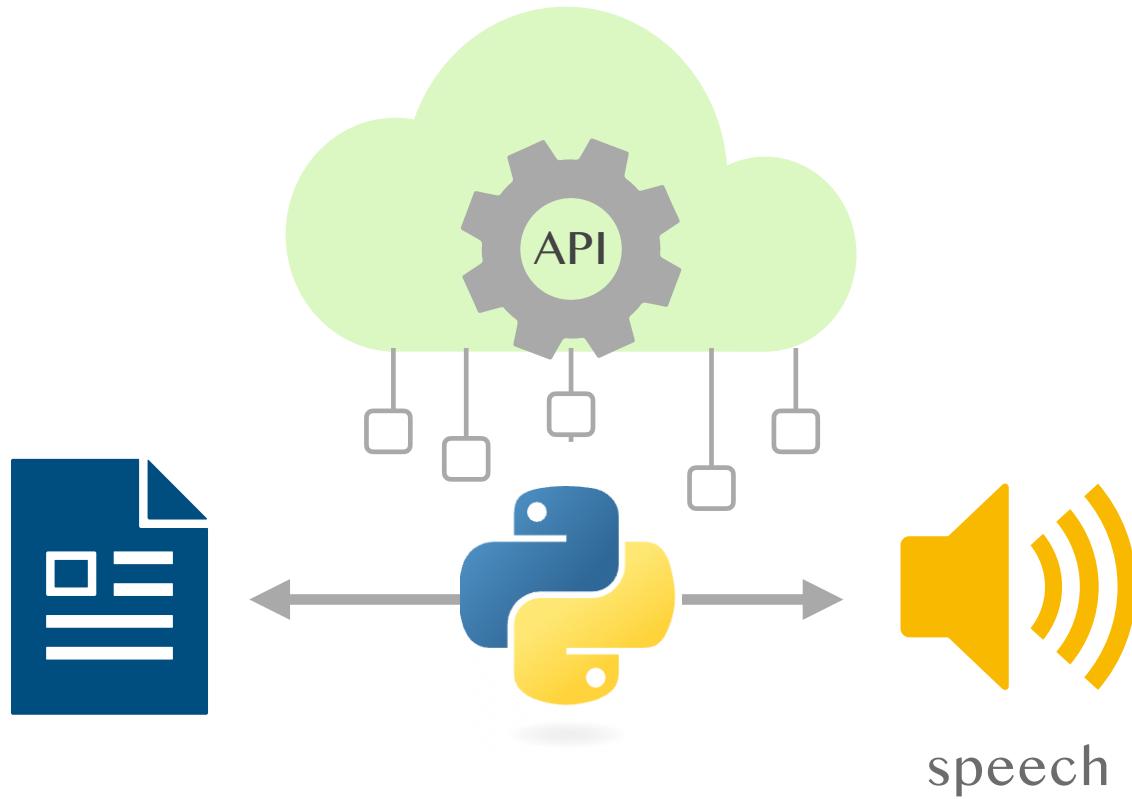
Output



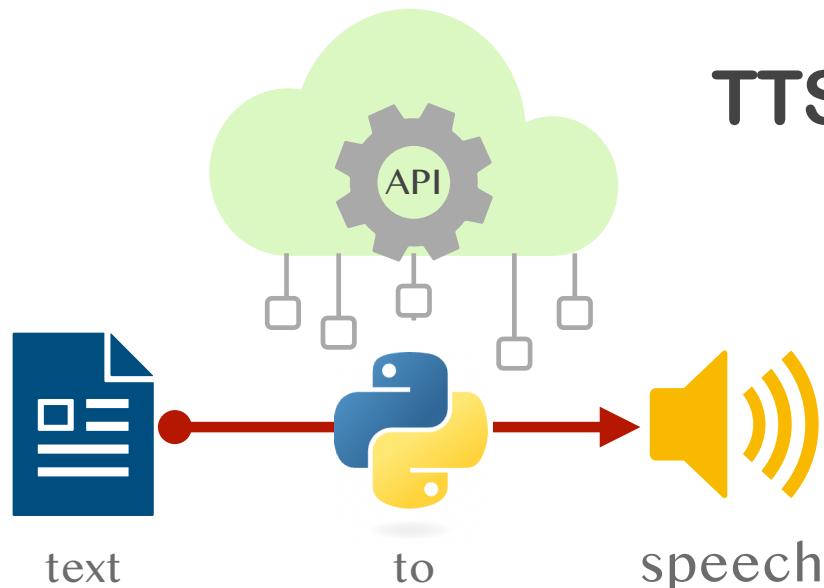
AI-Generated Art Won a Prize. Artists Aren't Happy. - The New York Times

Visit

API (Application Programming Interface)



TTS (Text-to-Speech)



OffiDocs

Voice Reader: Text to Speech

Freddy Rincón helped to lead the golden generation of Colombian football.

CNN — It's rare that a footballer touches a nation's heart in the way that Freddy Rincón did in Colombia.

That is why the 55-year-old's death was met with an outpouring of emotion in his home country.

"This is a huge loss for his family and friends, and at the same time is a regrettable loss for our sport, where we'll miss him and remember him with great care, appreciation, respect and admiration," the Colombian Football Federation said in a statement.

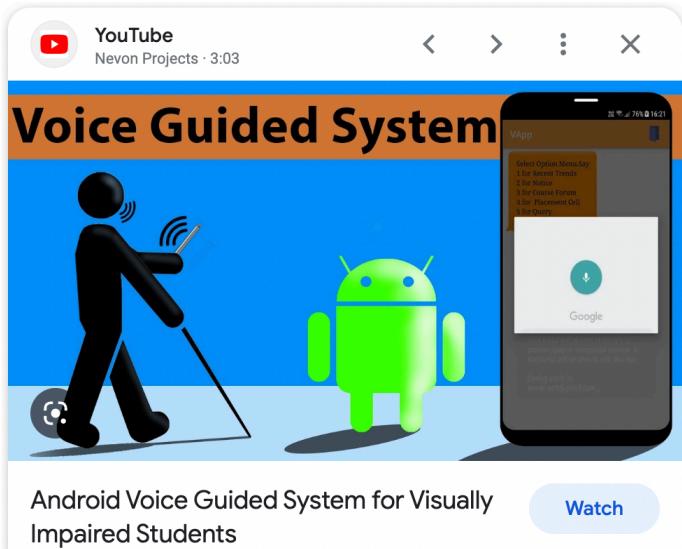
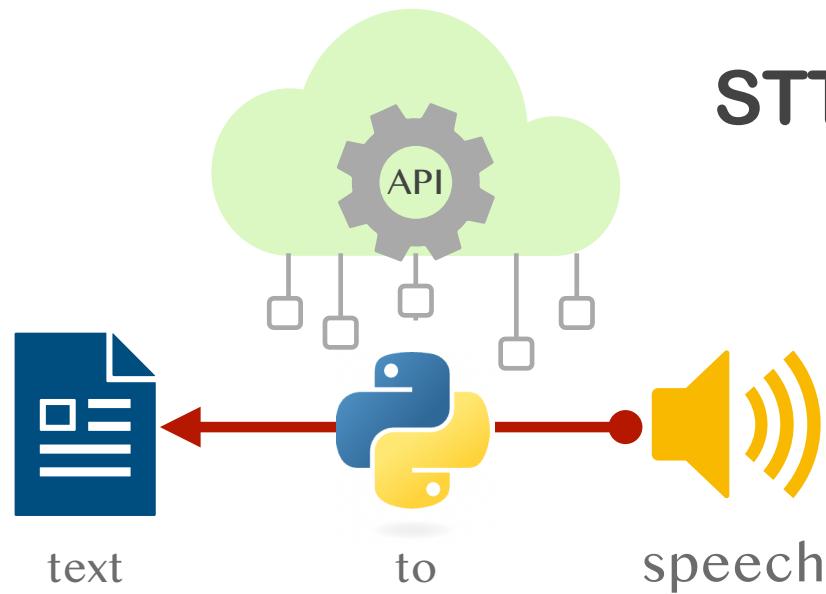
"We send strength, support and our condolences to his family, hoping that they can support this sad and painful loss."

An influential figure among Colombia's golden generation of players in the 1990s, the versatile midfielder signalled a new era for the country's football team.

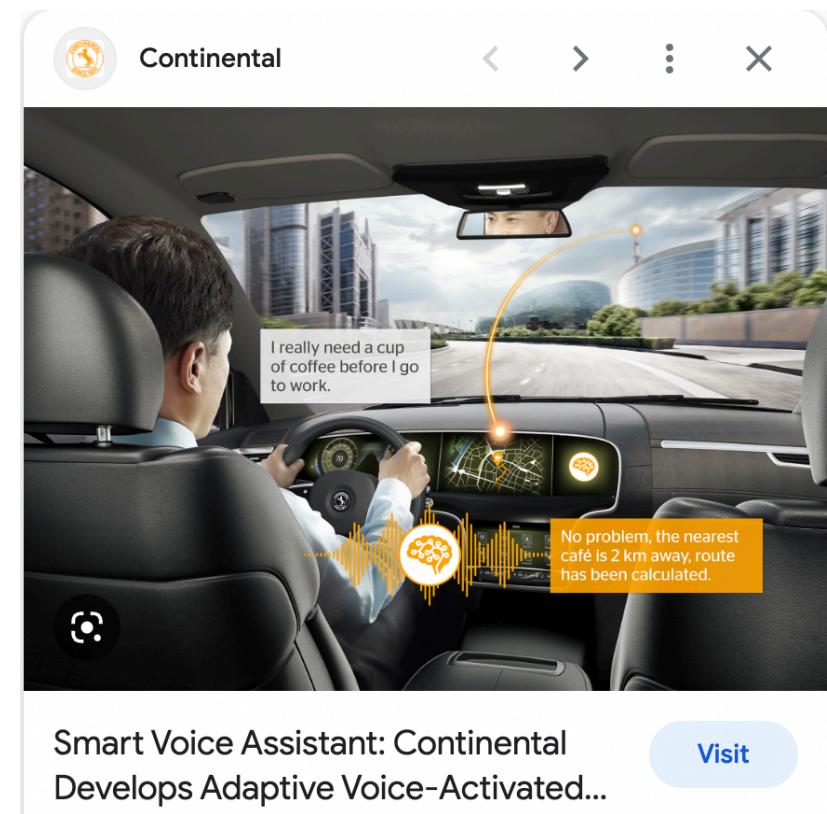
He played, and eventually captained, his country in 84 games, helping them to three World Cups, in 1990, 1994 and 1998.

Text to Speech: Voice Reader TTSSin Chrome with [Visit](#)

All images from Google open search



STT (Speech-to-Text)



All images from Google open search