

Signal Processing

Speech Signal Processing

Speech Information Processing

Role of language

- **Semantics:** The meanings of words, and relations among them.
- **Syntax:** The order of words, role of function words.
- **Phonology:** Individual phonemic segments, features, stressed and unstressed vowels.
- For example,
 - What is the phonemic inventory of English?
 - How does it function? The concept of contrast (e.g., **pat** vs. **bat**).
 - Why do we believe that it is psychologically real?
 - Why does the same phoneme give rise to different, acoustic realizations in different utterances? (e.g., In fluent speech, "Joe ate his soup" loses the /h/ of "his", and the /t/ of "ate" doesn't look like a /t/ in "Tom".)
 - What are the principles that lead to modifications of segments in different environments?
 - How are phonemes usually described in terms of features translated into phonetic representations? (e.g., /z/ is + voiced, /s/ is - voiced; same relation for many pairs, like f-v; patterning of sounds is beautifully captured by feature concept.)

New terms that you will know..

Voiced, Unvoiced, pitch, intensity, timbre, formants, speech production, vocal tract, vocal cord, phonemes, manner and place of articulation, co-articulation, linear prediction, homo-morphic filtering, spectrograms, speech coding, speech enhancement, speech recognition, acoustic modelling, time/pitch scale modification, speech synthesis, Human auditory system, speech perception, speech quality measures (MOS, PESQ).....

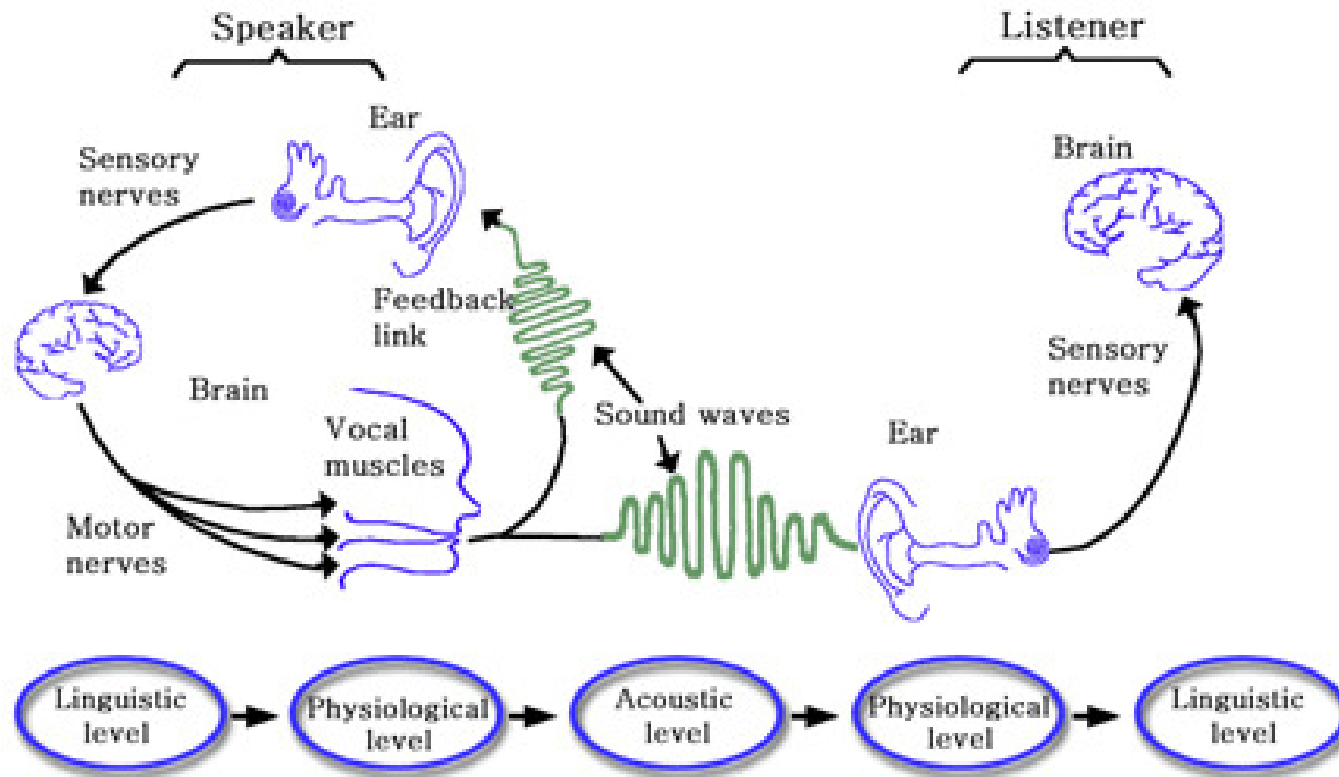
Speech Information Processing

**Speech Information Understanding and
Modeling**

What is information in speech and
how it is encoded?

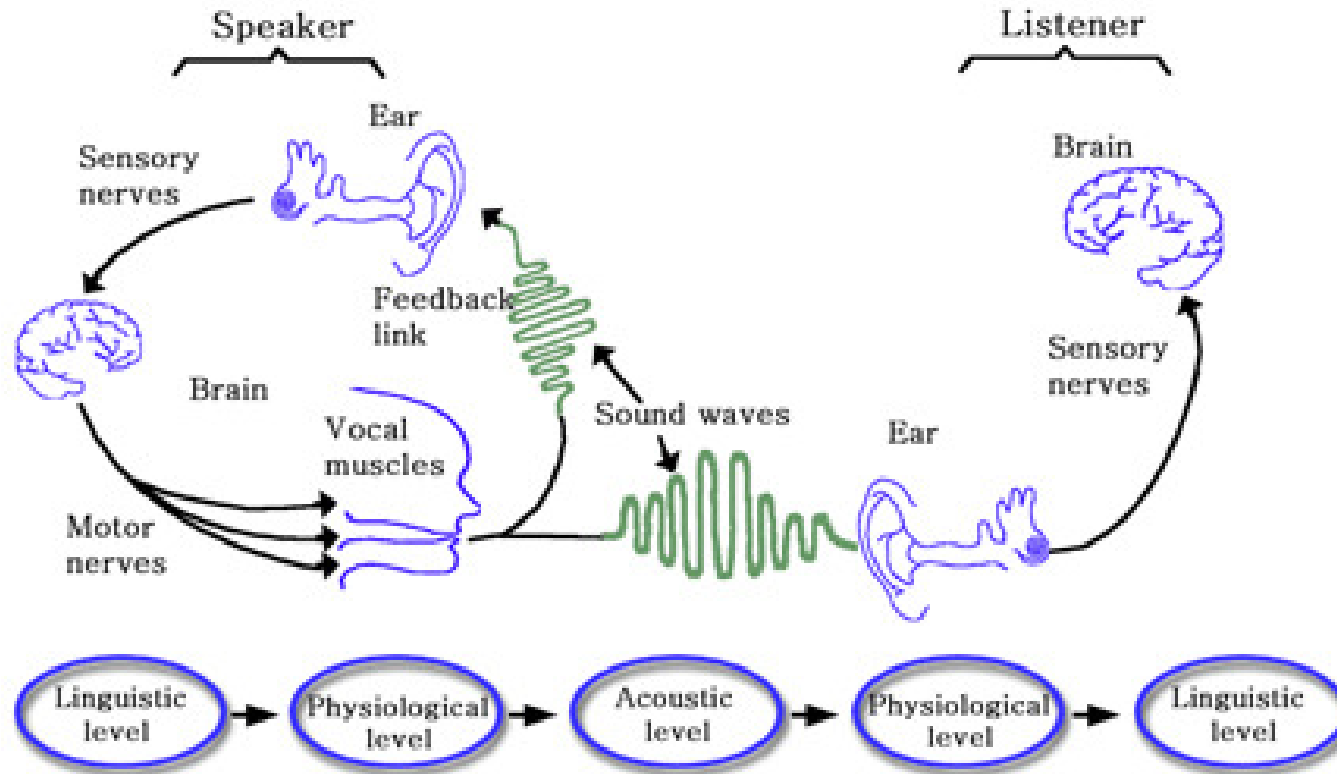
Lets give it a try

The Speech Chain



- Speech Research –
 - Speech Science
 - - Linguistics
 - - Physiology of Speech Production
 - - Acoustics
 - - Auditory Nervous System
 - - Psychophysics of Auditory System
 - - Cognitive Psychology
 - - Computer-based Algorithms
 - Speech Technology
 - Speech Recognition
 - Speaker Recognition
 - Speech Synthesis

The Speech Chain



← Speech Coding →

← Speech Synthesis →

← Speech Recognition →

← Speech Understanding →

← Speaker Recognition →

← Language Recognition →

- Challenges to machine speech processing:
 - Definition of information content
 - Multiple levels of information
 - Subjectivity of the listener
 - Robustness to
 - Interfering signals
 - Partial information
 - Algorithmic complexity

Recommended Readings

- J. L. Flanagan, *Speech Analysis, Synthesis, and Perception*, Springer-Verlag, 2nd Edition, Berlin, 1972
- J. D. Markel and A. H. Gray, Jr., *Linear Prediction of Speech*, Springer-Verlag, Berlin, 1976
- B. Gold and N. Morgan, *Speech and Audio Signal Processing*, J. Wiley and Sons, 2000
- J. Deller, Jr., J. G. Proakis, and J. Hansen, *Discrete – Time Processing of Speech Signals*, Macmillan Publishing, 1993
- D. O’Shaughnessy, *Speech Communication, Human and Machine*, Addison-Wesley, 1987
- S. Furui and M. Sondhi, *Advances in Speech Signal Processing*, Marcel Dekker Inc, NY, 1991
- R. W. Schafer and J. D. Markel, Editors, *Speech Analysis*, IEEE Press Selected Reprint Series, 1979
- D. G. Childers, *Speech Processing and Synthesis Toolboxes*, John Wiley and Sons, 1999
- K. Stevens, *Acoustic Phonetics*, MIT Press, 1998
- J. Benesty, M. M. Sondhi and Y. Huang, Editors, *Springer Handbook of Speech Processing and Speech Communication*, Springer, 2008.

Recommended Readings

Speech Coding:

- A. M. Kondoz, *Digital Speech: Coding for Low Bit Rate Communication Systems-2nd Edition, John Wiley and Sons, 2004*
- W. B. Kleijn and K. K. Paliwal, Editors, *Speech Coding and Synthesis, Elsevier, 1995*
- P. E. Papamichalis, *Practical Approaches to Speech Coding, Prentice Hall Inc, 1987*
- N. S. Jayant and P. Noll, *Digital Coding of Waveforms, Prentice Hall Inc, 1984*

Recommended Readings

Speech Synthesis:

- *T. Dutoit, An Introduction to Text - To-Speech Synthesis, Kluwer Academic Publishers, 1997*
- *P. Taylor, Text-to-Speech Synthesis, Cambridge University Press, 2008*
- *J. Allen, S. Hunnicutt, and D. Klatt, From Text to Speech, Cambridge University Press, 1987*
- *Y. Sagisaka, N. Campbell, and N. Higuchi, Computing Prosody, Springer Verlag, 1996*
- *J. VanSanten, R. W. Sproat, J. P. Olive and J. Hirschberg, Editors, Progress in Speech Synthesis, Springer Verlag, 1996*
- *J. P. Olive, A. Greenwood, and J. Coleman, Acoustics of American English, Springer Verlag, 1993*

Recommended Readings

Speech Recognition:

- *L. R. Rabiner and B. H. Juang, Fundamentals of Speech Recognition, Prentice Hall Inc, 1993*
- *X. Huang, A. Acero and H-W Hon, Spoken Language Processing, Prentice Hall Inc, 2000*
- *F. Jelinek, Statistical Methods for Speech Recognition, MIT Press, 1998*
- *H. A. Bourlard and N. Morgan, Connectionist Speech Recognition-A Hybrid Approach, Kluwer Academic Publishers, 1994*
- *C. H. Lee, F. K. Soong, and K. K. Paliwal, Editors, Automatic Speech and Speaker Recognition, Kluwer Academic Publisher, 1996*