LTI
Language, Technology and the Internet

New Mediums: Email; Usenet; Blogs, Chat, …

Francis Bond

Department of Asian Studies
Palacký University
https://fcbond.github.io/
bond@ieee.org

Lecture 4

Revision of Speech and Language Technology

➤ The need for speech representation

➤ Storing sound

➤ Transforming Speech

  ➤ Automatic Speech Recognition (ASR): sounds to text
  ➤ Text-to-Speech Synthesis (TTS): texts to sounds

➤ Speech technology — the Telephone!

How good are the systems in 2020?

English Tasks	WER%
LibriSpeech audiobooks 960hour clean	1.4
LibriSpeech audiobooks 960hour other	2.6
Switchboard telephone conversations between strangers	5.8
CALLHOME telephone conversations between family	11.0
Sociolinguistic interviews, CORAAL (AAL)	27.0
CHiMe5 dinner parties with body-worn microphones	47.9
CHiMe5 dinner parties with distant microphones	81.3

Chinese (Mandarin) Tasks	CER%
AISHELL-1 Mandarin read speech corpus	6.7
HKUST Mandarin Chinese telephone conversations	23.5

Rough Word Error Rates (WER = % of words misrecognized) reported around 2020 for ASR on various American English and Chinese recognition tasks.

Note, the error rate for transcribing conversations between humans is much higher; 5.8 to 11% for the Switchboard and CALLHOME corpora. (Jurafsky and Martin, 2023, Ch. 16, p4)

Why is ASR so difficult?

➢ Speaker variability

    ➢ Gender
    ➢ Dialect/Foreign Accent
    ➢ Individual Differences: Physical differences; Language differences (idiolect)

➢ Many, many rare events

    ➢ 300 out of 2000 diphones in the core set for the AT&T NextGen system occur only once in a 2-hour speech database

Prosody of Emotion

➤ Excitement: Fast, very high pitch, loud

➤ Hot anger: Fast, high pitch, strong, falling accent, loud

➤ Fear: Jitter

➤ Sarcasm: Prolonged accent, late peak

➤ Sad: Slow, low pitch

The main determinant of "naturalness" in speech synthesis is not "voice quality", but natural-sounding prosody (intonation and duration)

Richard Sproat

The Telephone

Speech like	Text like
time bound	space bound
spontaneous	contrived
face-to-face	visually decontextualized
loosely structured	elaborately structured
socially interactive	factually communicative
immediately revisable	repeatedly revisable
prosodically rich	graphically rich

Overview

A look at some different modalities of communication

➢ Email

➢ Usenet

➢ Chat

➢ Blogs

Email

Speech like	Text like
time bound*	space bound (deletable)
spontaneous*	contrived*
face-to-face	visually decontextualized
loosely structured*	elaborately structured*
socially interactive*	factually communicative
immediately revisable	repeatedly revisable*
prosodically rich	graphically rich *

How is it different from snailmail?

Email vs Mail

➤ Faster (potentially)

➤ Easy to include multiple participants

➤ Easy to include the original in the reply

➤ Copies stored with each participant
   (and temporarily at all intermediate nodes)

Email Nettiquette

http://pthree.org/2010/09/18/email-netiquette-part-1/

1. Use plain text (preferred) or HTML

2. Top-post only when forwarding. Bottom-post otherwise.

3. Trim your replies.

4. Keep you signature under five lines, and use the signature separator "-- " (dash, dash, space).

5. Do not attach unnecessary files, keep attachments small, and don't attach proprietary formats.

6. Keep the width of your message under 80 characters

7. Use a client that sends threading headers.

8. Reply only to the necessary people (don't abuse CC: or "reply to all").

9. Be short and concise. Don't ramble (stay on topic).

10. Break up your paragraphs.

11. Use proper spelling, grammar and punctuation (avoid CAPS).

12. Don't answer spam, and don't send out spam.

These already seem quaint to me (FCB) although I try and follow them.

A Gricean Analysis

Maxim of Quantity

➤ Make your contribution as informative as is required (for the current purposes of the exchange).
➤ Do not make your contribution more informative than is required.

Maxim of Quality

➤ Do not say what you believe to be false.
➤ Do not say that for which you lack proper evidence.

Maxim of Relation

➤ Be relevant.

Maxim of Manner

➤ Be perspicuous [= be easily understood]
➤ Avoid obscurity of expression.
➤ Avoid ambiguity
➤ Be brief (avoid unnecessary prolixity)
➤ Be orderly

Discuss each recommendation and try to see which Maxim or Maxims fit.

Top-Posting

Trim the message, leaving appropriate context, then
reply below.
> How should I reply to email then?
>> No.
>>> Should I include quotations after my reply?
>>>> Because it messes up the order in which people normally
>>>> read text.
>>>>> Why is top-posting such a bad thing?
>>>>>> Top-posting.
>>>>>>> What is the most annoying thing in email?

Bottom-Posting

```
>>>>>>> What is the most annoying thing in email?
>>>>>> Top-posting.
>>>>> Why is top-posting such a bad thing?
>>>> Because it messes up the order in which people normally
>>>> read text.
>>> Should I include quotations after my reply?
>> No.
> How should I reply to email then?
Trim the message, leaving appropriate context, then
reply below.
```

```
[snip]
> How should I reply to email then?
Trim the message, leaving appropriate context, then
reply below.
```

Email Salutations

Email Endings

Netspeak

Internet specific slang/jargon

There was a time when only computer scientists/researchers had email and access to the internet, …

➤ Inspired by shared background

➤ `<rant>`I can't stand this`</rant>`
➤ I hate^H^H^H^Hlove that idea   (`^H` is backspace on a vt100)
➤ lusers (users as seen by Systems Administrators)
➤ suits

➤ Inspired by text limitations

➤ `GREAT;*great*;\great/;` g r e a t
➤ `:-)   (^_^)   (;o;)` ‿

➤ Telcos and then Unicode have added many emoji, somewhat depressing this exuberance

---

Usenet (asynchronous)

Usenet "User's Network" is a worldwide distributed discussion system. It was developed from the general-purpose Unix-to-Unix Copy (UUCP) dial-up network architecture. Users read and post messages (called articles or posts, and collectively termed news) to one or more topic categories, known as newsgroups. Usenet

Speech like	Text like
time bound*	space bound
spontaneous*	contrived*
face-to-face	visually decontextualized
loosely structured*	elaborately structured
socially interactive*	factually communicative
immediately revisable	repeatedly revisable
prosodically rich	graphically rich

Similar to: Bulletin boards; the WELL; some chat; Mailing Lists; forums

➤ `linguist-list`

➤ `sci.lang.japan` FAQ: `http://www.sljfaq.org/afaq/afaq.html`

Netspeak

➤ Need to be in-group:

cow orker Coworker

clue "You couldn't get a clue during the clue mating season in a field full of horny clues if you smeared your body with clue musk and did the clue mating dance." Edward Flaherty (`talk.bizaare`)

➤ Posting (top/middle/bottom, quoting and trimming)

➤ Appropriate length

l33tspeak

➤ Designed to confuse n00bs

➤ Full of internet memes

   ➤ "All your base are belong to us"
   ➤ Owned and Pwn3d

➤ Pr0n

```
 -------------------------------------------------------------------
|                                                                   |
|          |                          \ | /                     |       |
|          |\                          \|/                      /|      |
|     |XXXXX||>>>>>>>>>>>>>>>>>>>>>>>>  ((*))   <<<<<<<<<<<<<<<<<<<<<<<<||XXXXX|  |
|          |/                          /|\                      \|      |
|          |                          / | \                      |       |
|                                                                   |
| Death Star                                                        |
| War Lord of the West                            Dkline@doc.bmd.trw.com |
|_____|
```

The original bad signature.

--------------------------- ALT.FAN.WARLORD FORM ---------------------------

YOUR SIGNATURE IS:

___ WIDER THAN 80 COLUMNS LONG OR LONGER. PLEASE ADJUST.
___ OVER 4 LINES, THE NET STANDARD.
___ MORE THAN TWICE THE NET STANDARD OF 4 LINES IN SIZE.
___ FAR LARGER THAN ANY WORTHWHILE COMMENTARY OF WHICH YOU MAY BE CAPABLE.

___ YOU HAVE ___ SNAIL MAIL ADDRESSES WHICH NO ONE REALLY USES.
___ YOU HAVE ___ PHONE NUMBERS FOR PEOPLE TO USE IN PRANK CALLS.
___ YOU HAVE ___ BEEPER NUMBERS FOR PEOPLE TO USE IN PRANK CALLS.
___ YOU HAVE LISTED MORE THAN ONE EMAIL ADDRESS.
    ARE YOU REALLY EXPECTING PEOPLE TO SEND EMAIL TO YOU IN TRIPLICATE?

YOUR SIGNATURE CONTAINS :

___ RUSH QUOTE.
___ UGLY BORDER.

```
___   LAME LOGIN NAME.
___   STUPID SELF-QUOTE.
___   OBNOXIOUS UNREADABLE FONT.
___   STUPIDLY LARGE DISCLAIMER.
___   TOO MUCH WHITE SPACE. CAN THIS BE REDUCED TO REACH THE 4 LINE LIMIT?
___   LAME ASCII GRAPHIC OF (Choose all that apply):
      __  AMIGA CHECK
      __  AUSTRALIA
      __  BICYCLE
      __  COMPANY LOGO
      __  SWORD
      __  USS ENTERPRISE


COMMENTS:
___   YOU HAVE GREATLY MISUNDERSTOOD THE PURPOSE OF A SIGNATURE FILE.
___   YOU HAVE GREATLY MISUNDERSTOOD THE PURPOSE OF ALT.FAN.WARLORD
___   YOU HAVE GREATLY MISUNDERSTOOD THE PURPOSE OF USENET.
___   YOU MUST HAVE SPENT YOUR ENTIRE LIFE IN A SKINNER BOX
      TO BE THIS CLUELESS.


___   THIS HAS BEEN WARLORDED BEFORE.
```

A different approach

NANYANG TECHNOLOGICAL UNIVERSITY

Daniel LEE (Assoc Prof) | Director | Computer Research Centre | Nanyang Technological University
NX.X-BX-XX, 50 Nanyang Avenue, Singapore 639798
Tel: (65) 6XXX-XXXX GMT+8h | Fax: (65) XXXX-XXXX | Email: yyy@ntu.edu.sg | Web: www.ntu.edu.sg

CONFIDENTIALITY: This email is intended solely for the person(s) named. The contents may be confidential and/or privileged. If you are not the intended recipient, please delete it, notify us, and do not copy or use it, nor disclose its contents. Thank you.

Towards A Sustainable Earth: Print Only When Necessary

The corporate email signature was launched on 6 October 2009 and is designed to project a distinct NTU identity and to achieve consistency in our communications, in particular with external parties. The email signature will also present essential information, such as contact details, for the convenience of those with whom the staff is communicating.

➢ What purpose does the disclaimer serve?

➢ Do we need all the contact info all the time? and the NTU logo?

Chat (synchronous)

Speech like	Text like
time bound*	space bound
spontaneous*	contrived
face-to-face	visually decontextualized
loosely structured*	elaborately structured
socially interactive*	factually communicative
immediately revisable	repeatedly revisable
prosodically rich	graphically rich

IRC; MSN;

Netspeak

➤ Need to use few key strokes:

imho In My Humble Opinion
rtfm Read The Fine Manual
brb Be Right Back
btw By The Way
cya See You

➤ Lower case

➤ Lack of apostrophes (they are hard on phone keyboards)

➤ Text input is slower than speech

➤ Especially on a ten key (although it is getting faster)

Very weird discourse

➢ Multiple Participants

➢ All can be heard/read

➢ Often lagging and overlapping

   ➢ Typically break down into several sub-conversations

Blogs

Speech like	Text like
time bound	space bound
spontaneous*	contrived*
face-to-face	visually decontextualized
loosely structured*	elaborately structured*
socially interactive^{comments}	factually communicative
immediately revisable	repeatedly revisable*
prosodically rich	graphically rich *

Combination of diary and broadsheet

Moderation

Teresa Nielsen Hayden (2005) *Virtual panel participation*

Some things I know about moderating conversations in virtual space:

1. There can be no ongoing discourse without some degree of moderation, if only to kill off the hardcore trolls. It takes rather more moderation than that to create a complex, nuanced, civil discourse. If you want that to happen, you have to give of yourself. Providing the space but not tending the conversation is like expecting that your front yard will automatically turn itself into a garden.

2. Once you have a well-established online conversation space, with enough regulars to explain the local mores to newcomers, they'll do a lot of the policing themselves.

3. You own the space. You host the conversation. You don't own the community. Respect their needs. For instance, if you're going away for a while, don't shut down your comment area. Give them an open thread to play with, so they'll still be there when you get back.

4. Message persistence rewards people who write good comments.

5. Over-specific rules are an invitation to people who get off on gaming the system.

6. Civil speech and impassioned speech are not opposed and mutually exclusive sets. Being interesting trumps any amount of conventional politeness.

7. Things to cherish: Your regulars. A sense of community. Real expertise. Genuine engagement with the subject under discussion. Outstanding performances. Helping others. Cooperation in maintenance of a good conversation. Taking the time to teach newbies the ropes. All these things should be rewarded with your attention and praise. And if you get a particularly good comment, consider adding it to the original post.

8. Grant more lenience to participants who are only part-time jerks, as long as they're valuable the rest of the time.

9. If you judge that a post is offensive, upsetting, or just plain unpleasant, it's important to get rid of it, or at least make it hard to read. Do it as quickly as possible. There's

no more useless advice than to tell people to just ignore such things. We can't. We automatically read what falls under our eyes.

10. Another important rule: You can let one jeering, unpleasant jerk hang around for a while, but the minute you get two or more of them egging each other on, they both have to go, and all their recent messages with them. There are others like them prowling the net, looking for just that kind of situation. More of them will turn up, and they'll encourage each other to behave more and more outrageously. Kill them quickly and have no regrets.

11. You can't automate intelligence. In theory, systems like Slashdot's ought to work better than they do. Maintaining a conversation is a task for human beings.

12. Disemvowelling works. Consider it.

13. If someone you've disemvowelled comes back and behaves, forgive and forget their earlier gaffes. You're acting in the service of civility, not abstract justice.

Disemvowelling

1

➢ Removing the vowels from an (obnoxious) comment

➢ On 30 October 2008, TIME magazine listed disemvoweling as #42 of their "Top 50 Inventions of 2008".

➢ Xeni Jardin, co-editor of Boing Boing, said of the practice, "the dialogue stays, but the misanthrope looks ridiculous, and the emotional sting is neutralized."

➢ Boing Boing producers claim that disemvoweling sends a clear message to internet forums as to types of behavior that are unacceptable.

➢ Try it here: `http://textmechanic.com/Disemvowel-Tool.html`

---

[1] `http://content.time.com/time/specials/packages/article/0,28804,1852747_1854195_1854185,00.html`

Licensed to disemvowel

No one's obliged to listen to you if you're being a jerk

KNOW ALL YE BY THESE PRESENTS THAT
TERESA NIELSEN HAYDEN
GRANTS THIS MODERATOR OFFICIAL PERMISSION TO ADVISE, CURB, ADMONISH, SORT OUT, LOOK ASKANCE AT, CHIDE, RESTRAIN, SIT ON, DISAPPROVE OF, REPRIMAND, DISCIPLINE, CALL TO ACCOUNT, REBUKE, CENSURE, UPBRAID, SHOOT DOWN, CHEW OUT, TELL OFF, GIVE NOTICE TO, DISEMVOWEL, DELETE, SUSPEND, PROSCRIBE, OR PERMANENTLY BAN TROLLS, KOOKS, SOCKPUPPETEERS, MONOMANIACS, BULLIES, AND OTHER
ILL-NATURED AND VEXATIOUS PERSONS.

ISSUED TO: _____

Denizens of the Net

Troll someone who posts inflammatory, extraneous, or off-topic messages in an online community with the primary intent of provoking other users into an emotional response or of otherwise disrupting normal on-topic discussion

Sock Puppet an online identity used for purposes of deception within an online community, for example, a fake online identity to praise, defend or create the illusion of support for one's self, allies or company.

Meat Puppet someone recruited to an online space by an existing member to support their position.

Time to pick your medium

➢ Choose something interesting, that you know about, or want to know about

➢ Take a look at similar papers from the past
  Papers from earlier years

➢ Try to add something new, ...

Acknowledgments

➤ Much Data taken from Crystal (2006, ch 2,4–5,8)

*

References

Crystal, D. (2006). *Language and the Internet*. Cambridge University Press, 2nd edition.

Jurafsky, D. and Martin, J. H. (2023). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. Prentice Hall, 3rd edition. (draft).