UNIVERSITÉ BISHOP'S UNIVERSITY

CS 469 / CS 569: Special Topics in Computer Science: Human-Computer Interaction

Expressive Human and Command Languages

Dr. Mohammed Ayoub Alaoui Mhamdi
Bishop's University
Sherbrooke, Qc, Canada
malaoui@ubishops.ca

Expressive Human and Command Languages

Topics

1. Introduction
2. Speech recognition
3. Speech production
4. Human language technology
5. Traditional command languages Introduction

Introduction

- The dream of speaking to computers and having computers speak has long lured researchers and visionaries
- Arthur C. Clarke's 1968 fantasy of the HAL 9000 computer in the book and movie 2001: A Space Odyssey has set the standard for performance of computers in science fiction and for developers of natural language systems
- The reality is more complex

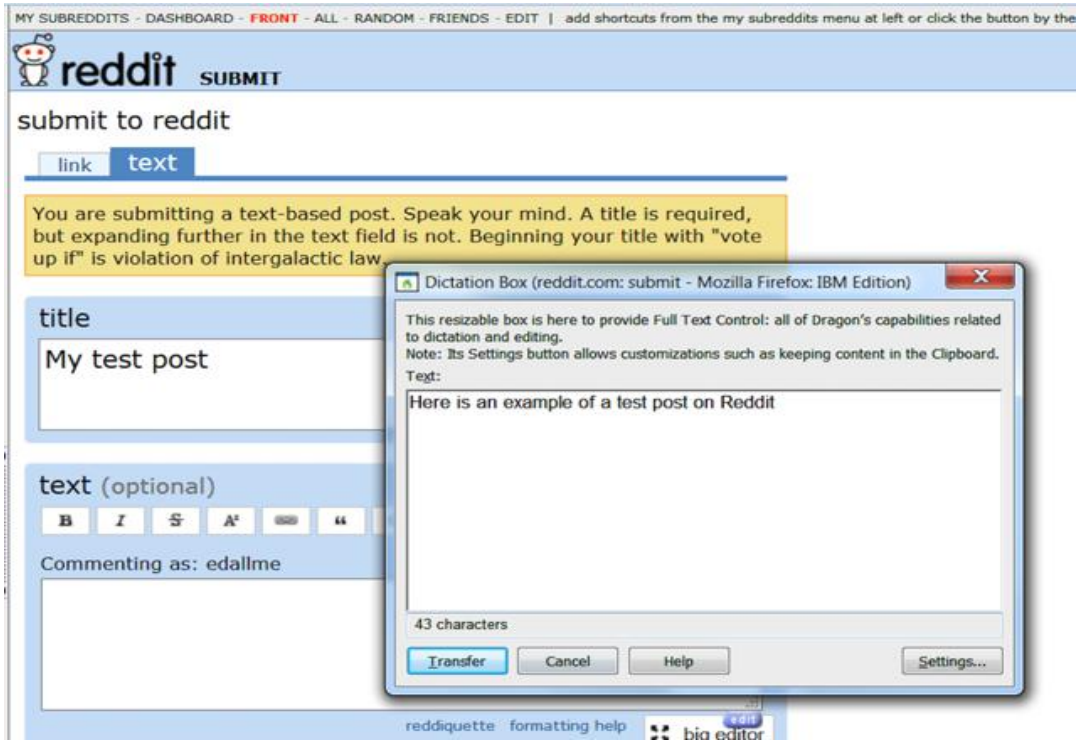
Speech Technologies

- Store and replay (museum guides)
- Dictation (document preparation, web search)
- Close captioning, transcription
- Transactions over the phone
- Personal "assistant" (common tasks on mobile devices)
- Hands-free interaction with a device
- Adaptive technology for users with disabilities
- Translation
- Alerts
- Speaker identification

Speech Recognition

- The place for spoken interaction
- Speech recognition applications
- Designing spoken interaction
- Spoken prompts and commands

Spoken Interaction



- Using Nuance Dragon™ speech dictation and a head mouse (as made visible by the little silver dot on his forehead), a computer scientist is able to overcome a temporary hand disability (http://www.nuance.com/dragon/index.htm)

Speech recognition and generation: opportunities

- When users have physical impairments
- When the speaker's hands are busy
- When mobility is required
- When the speaker's eyes are occupied
- When harsh or cramped conditions preclude use of a keyboard
- When application domain vocabulary and tasks is limited
- When the user is unable to read or write (e.g. children)

Speech recognition and generation: obstacles to speech recognition

- Interference from noisy environments and poor-quality microphones
- Commands need to be learned and remembered
- Recognition may be challenged by strong accents or unusual vocabulary
- Talking is not always acceptable (e.g. in shared office, during meetings)
- Error correction can be time consuming
- Increased cognitive load compared to typing or pointing
- Math or programming difficult without extreme customization

Speech recognition and generation: obstacles to speech production

- Slow pace of speech output when compared to visual displays
- Ephemeral nature of speech
- Not socially acceptable in public spaces (also privacy issues)
- Difficulty in scanning/searching spoken messages

Voice-activated Digital Assistants

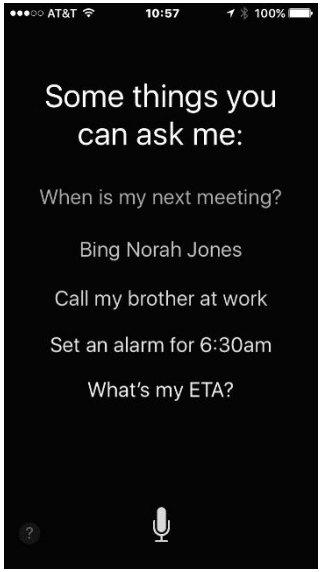
- A few years ago, you would only see someone talking into their phone if somebody was on the other side
- Fast forward a bit and now talking to your phone when you are not on a call is no big deal
- Siri for iPhone revolutionized the behavior, and nowadays it is common to see people use their voice to control their phones

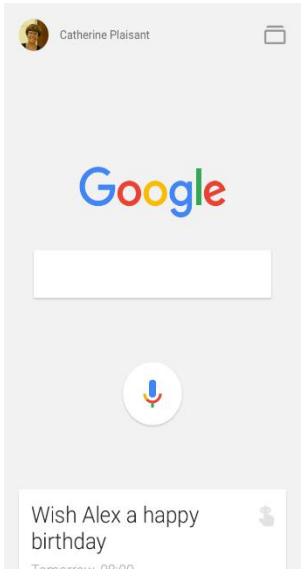
Designing spoken interaction (1 of 5)

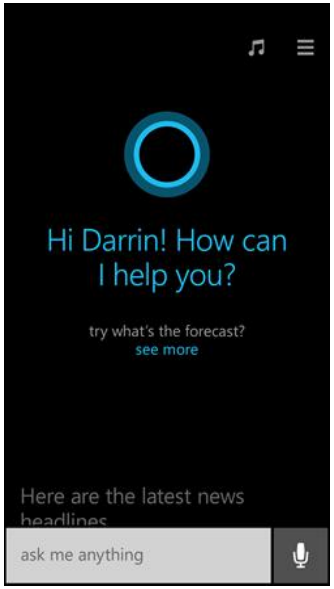
- Initiation
- Knowing what to say
- Recognition errors
- Correcting errors
- Mapping to possible actions
- Feedback and dialogs

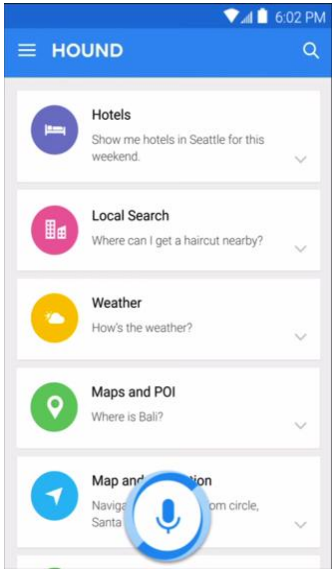
Designing spoken interaction
(2 of 5)

- Mobile devices assistants (from left to right: Siri, GoogleNow, Cortana and Hound) all have similar microphone buttons, but different ways of presenting suggestions



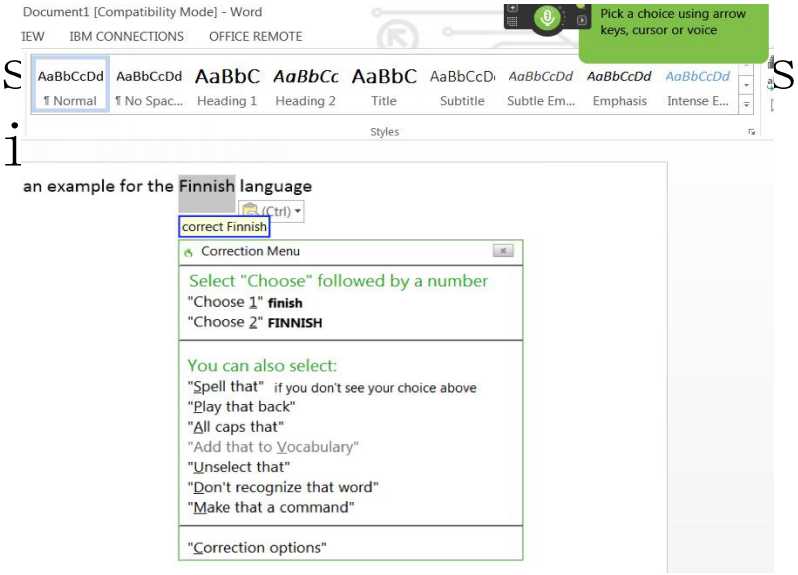




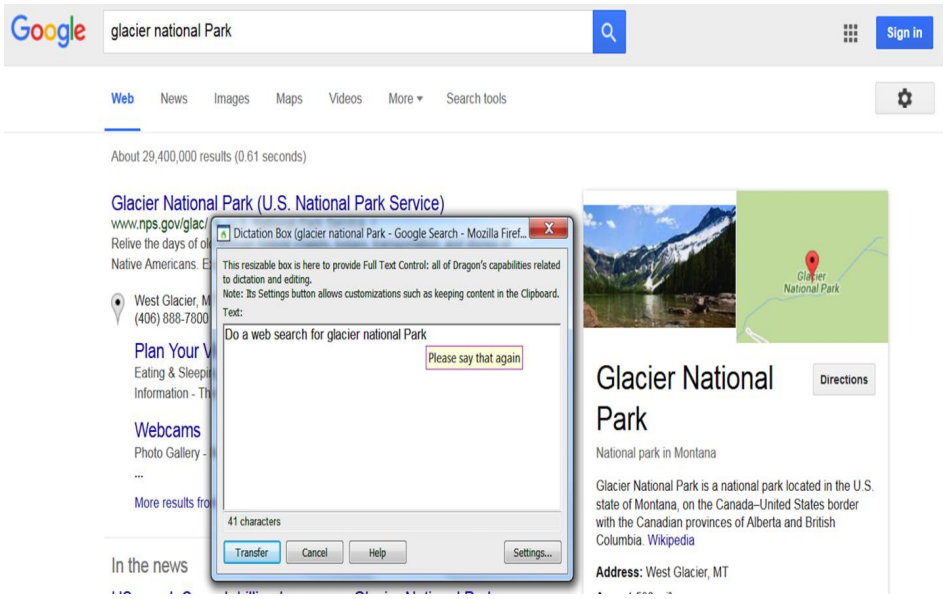


Designing spoken interaction (3 of 5)

- Correcting a word during dictation using Nuance Dragon™.
- After saying “Correct finnish” the word is selected and possible corrections are displayed in a menu, along with additional commands such as “Spell that”
- Users can us s, or voice to specify thei



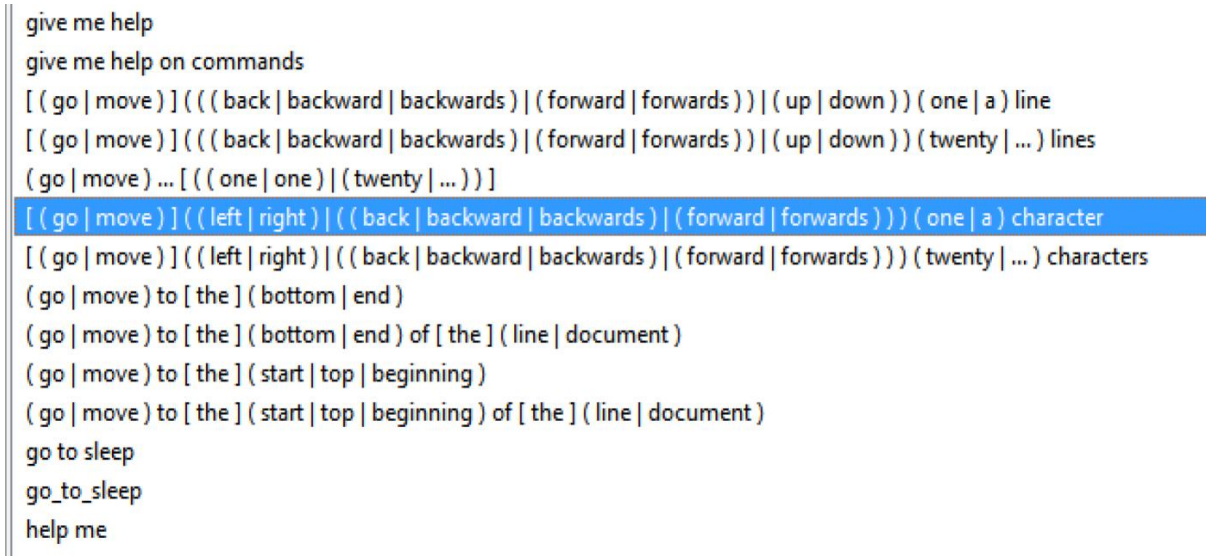
Designing spoken interaction (4 of 5)



- It can be difficult to remember what exact command will accomplish the task
- In this example when the user said “Search the web for Glacier National Park” a Google search was launched and a search executed with the correct terms, but when the user said “Do a web search for Glacier National Park” the text was indeed accurately recognized but not as a command, so the text was placed in the Nuance Dragon™ dictation box

Designing spoken interaction (5 of 5)

- A small subset of the rich set of commands used in the Nuance Dragon™ speech recognition system
- Synonyms are included and used consistently



Speech Production (1 of 2)

- Speech production is usually successful when the messages are simple and short; and users' visual channels are overloaded
- There are three general methods to produce speech:

1. Formant synthesis – machine-generated speech using algorithms
2. Concatenated synthesis – uses tiny, recorded human speech segments
3. Canned speech – fixed, digitized speech segments

Speech Production (2 of 2)

- Examples:
  - Audio books or audio tours
  - Instructional systems
  - Online help systems
  - Alerts and warnings
  - Applications for the visually impaired

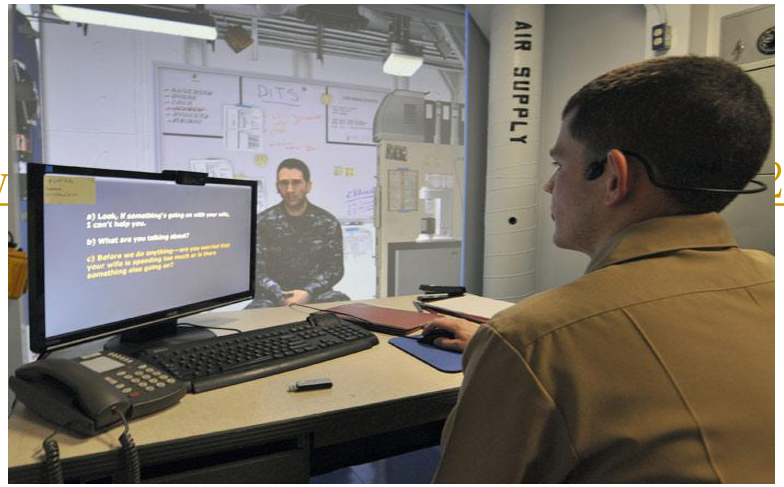
Human Language Technology (1 of 3)

- Machines that understand natural language
- Natural language interaction (NLI)
  - Series of exchanges or "dialog" is difficult to design and build, on even a single topic
  - Current successes often rely on statistical methods based on the analysis of vast textual or spoken data from millions of users
- Example applications and methods include:
  - Question answering strategies
  - Extraction and tagging, e.g. gathering data from a database of medical records
  - Human language text generation
  - Instructional systems
  - Language translators, e.g. Google Translate

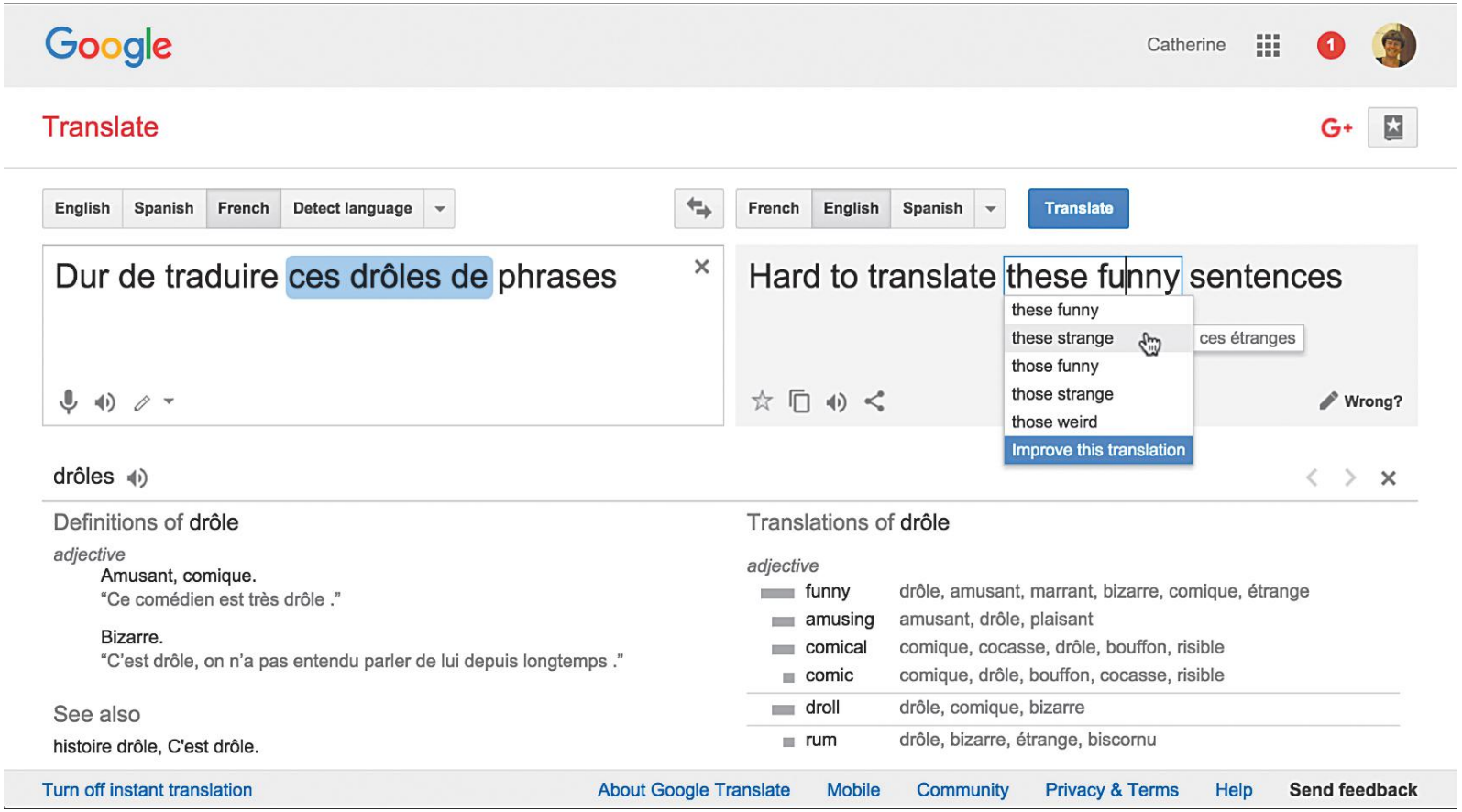
Human Language Technology (2 of 3)

- Using the Immersive Naval Officer Training System (INOTS) new navy officers can practice their counseling skills in a virtual reality environment
- Officers listen to an avatar and respond using spoken language, loosely following suggestions from multi-choice prompts presented on the screen and designed to match the learning objectives
- The interaction is constrained but assessment is facilitated

(Dyke, 2013; www.netc.navy 24 2.asp)



Human Language Technology (3 of 3)



Google Translate, showing a French sentence translated in English

Command Languages (1 of 2)

- Command languages are often preferred by expert users who do not want to drag and drop items for repeated steps.
- A command language example is the Unix command used to delete blank lines from a file
  - grep -v ^$ filea > fileb
- Casual users favor GUIs but both styles of interface can be made available successfully
- Other examples that behave like command languages:
  - Web addresses (URLs) can be seen as a form of command language
  - Twitter addresses
  - Database query languages

Command Languages (2 of 2)

- Using the Sublime text editor a user is doing a search and replace in a data table using regular expressions
- Typing " \t.*? Police" in the search box searches for a tab followed by zero or more character, a space, and then by "Police"
- The patterns found in the document are highlighted with a thin black line in the document, showing that both "local police" and "state police" have been found and selected
- An overview of the entire document is visible on the right, revealing the