Lecture 1

The Study of Language

Introduction

- *Linguistics* is the study of language in scientific form

    - *Language* overall is a complex system of communication (written, spoken, etc.)
    - "A language" is an example of such a system in practice

- **Language is a uniquely human ability**

- Other species can communicate but the creative and complex capacity of humans to synthesize new ideas is unparallelled

- Languages are widely varying but a key aspect of identity and culture:

    - Same language, different regions
    - Same language, different cultures
    - Same language, different social situations
    - Different languages altogether

- Why study language? It helps in many ways:

    - Learn to develop better tools for language instruction
    - Design better computers that can interact with humans
    - Treat people with speech/language disorders more effectively
    - Help provide the appropriate toolkit for learning a language more easily
    - Aid in human communication through *interpretation* and *translation* of languages

- Human language is an extremely complex phenomenon, so hard to study all aspects of it

    - *Linguists* work on researching different aspects of human languages
    - Key goal is to discover how human language works

- Scientific study of language:

    - Language is systematic, from individual sounds to entire discourses
    - An infinite number of ideas can be expressed and re-expressed in languages
    - Despite variance, languages are strangely universal
        - Some characteristics are shared by all languages...
        - ...others are unique to a single language
    - Languages often have arbitrary properties that don't follow systematic rules
    - Children can acquire language without being explicitly taught (*innate* ability)
    - Languages change over time, even if speakers do not want them to

- All humans master at least ONE language

    - Some aspects of language appear to be innate, which is why they are so universal

Linguistic Competence

- Native English speakers (or other language speakers) know a great deal about their language

  - However, native knowledge ≠ technical knowledge
  - May not be able to accurately describe how the language works
  - "Feel" the language for what is right as opposed to conclusions based on formal study

- However, being a native speaker does not translate to awareness (i.e. "hidden knowledge")

  - This UNCONSCIOUS ability to effectively use a language is *linguistic competence*

- Linguistic competence vs. *linguistic performance*

  - Competence: Unseen potential to speak or write a language
  - Performance: Observable realization of potential
    - i.e. what is actually done with the linguistic competence in practice
    - However, revelation of potential is NOT the same as conscious awareness of how it works

- Could have *performance errors*: inability to remember a word, accidental mispronunciation, stumbling, *spoonerisms*

  - Reasons for errors: feeling tired/distracted, particularly difficult utterances, etc.
  - However, errors do not detract from linguistic competence

- Linguistic competence cannot be observed directly

  - Performance serves as basis for making hypotheses and drawing conclusions
  - Linguists must disregard imperfections in performance like occasional errors

- **Language is used for communication of ideas**:

  - i.e. *mind-to-mind* communication
  - Communication systems can be quite diverse and don't just have to be written/spoken, e.g.:
    - Honking a car horn to indicate that someone should move
    - Using semaphore flags for signalling purposes
  - Key idea: end-to-end *communication chain*
    - *Transmitter* (sometimes called *host*): source of information
    - *Receiver*: destination of information
    - A *signal* is transferred from the transmitter to the receiver
    - Both parties must be competent in their own language for this to work

- Communication chain (we will study each step in time):

  1. Transmitter thinks of idea to communicate
  2. Transmitter picks out words to express the idea
  3. Transmitter puts these words together in a certain order, following language rules
  4. Transmitter figures out how to pronounce or write these words down
  5. Transmitter sends those pronunciations or writing cues to their vocal or kinesthetic anatomy
  6. Transmitter speaks or writes, sending signals through the air or on paper
  7. Receiver perceives the sounds or written words as they hear/see it
  8. Receiver decodes the sounds or written words by interpreting them as language
  9. Receiver receives the communicated idea by connecting it with knowledge from brain
  10. Receiver may respond, thereby starting the communication chain over again

- The host and the receiver may not be able to communicate in the same language

  - In this case, need a *interpreter* and/or a *translator*
  - Interpreters/translators act as the bridge between Steps 6 and 7
  - Interpreters/translators (mostly) follow the communication chain themselves, except idea is already formed!
  - We will study translation in greater detail later

- Sometimes ideas could be lost in transmission or even in translation!

  - *Noise*: interference in the communication chain

Language Knowledge in Theory

- Consider the case of a natively-used language by individuals:

  - Spoken language (*natural language*): basic units used are speech sounds (*phonemes*)
  - Signalled language (*sign language*): analogous "small units" are used (*cheremes*)
  - Either way, native speakers innately know valid and invalid sounds
    - e.g. dog barking or door slamming CANNOT be confused as language
    - Can recognize which sounds are part of native language and which are not
    - Can recognize AND (re)produce these basic sounds without understanding the underlying mechanics

- *Phonetics*: study of sound of human speech

  - e.g. innately knowing the difference between the vowels in `bat`, `beat`, `boot`
  - Key: pronunciations may differ but knowledge is identical!

- Other knowledge is also present, however:

  - Knowing how the sounds work together in a systematic way to produce competent vocal strings
  - Knowing which sequences of sounds are valid and which ones are not
    - c.f. `pterodactyl`, `Ptolemy`, `captive`
    - The /pt/ sound is not present at beginnings of words in English (but /p/ is fine)
    - Other languages may allow /pt/ at beginnings of words (e.g. Greek)
    - Will explain the /·/ notation later!

- *Phonology*: study of systematic organization of sounds in languages

  - Distribution of speech sounds, especially in recognizing validity/invalidity
  - e.g. can innately predict that `spaff` is a possibly valid word but `fsap` is not, despite not knowing meanings
  - Phonological knowledge allows recognition of sounds and words spoken by different speakers
    - i.e. different pronunciations can be adjusted for without additional knowledge
    - Even regionally, people may not pronounce the same words the same way, after all

- *Speech* is generally thought of as a continuous stream of sound

  - Few pauses, and native speakers can usually analyze quick sequences without problems
  - e.g. `thedogisplayinginthebackyard` vs `the dog is playing in the backyard`
  - We can still pick out the differences by *parsing* (splicing apart) the sounds correctly

- Another key ability is to break down words into smaller, meaningful parts

  - e.g. can break apart `unbelievability` into `un-`, `believe`, `-able`, `-ity`
  - OTOH `ungiraffelike` is not a real word but it is easy to guess its meaning if it were
  - Can deduce that `nicely` is a word but `bookly` is not (`bookish` IS, though)

- *Morphology*: study of formation of words and their relationships with each other

  - Structure of words is important to study
  - e.g. stems, roots, prefixes, suffixes, parts of speech, intonation, stress
  - Context is very important for morphological study of language

- Internal knowledge of how words precisely combine to form phrases and sentences is also present:

  - Ability to construct/use sentences that brain has never seen before
  - Ability to pick apart *grammatical* constructs from *ungrammatical* constructs
    - Grammatical: `I will pick the package up at eight o'clock.`
    - Ungrammatical: `Package up pick at o'clock will the eight I.`
    - A lot of ungrammatical constructs still carry their original meaning, however!
    - A seasoned native would probably still be able to understand the second sentence above

- Sometimes it may be a little harder but the brain still is able to do it:
  - Grammatical: `I have a cup of pebbles` or `I have a cup of gravel`
  - Ungrammatical: `I have a cup of pebble` or `I have a cup of gravels`
  - Understanding the correct singular vs. plural usage is somewhat innate, but even natives make mistakes
  - (May be a property of English, which is a bit more flexible when it comes to meaning)

- *Syntax*: set of rules, principles, and processes that govern the structure of words in a language

  - e.g. word order, punctuation, grammar, logic, etc.
  - Syntax-oriented linguists focus on the discovery of *syntactic rules* used in formal languages
  - Can be a mathematically-based study, as shown in the study of the theory of computation in the computer sciences

- Assigning meaning to well-formed sentences is another key function of language communication

  - i.e. interpreting the meaning of sentences after parsing them correctly
  - Native speakers assign meanings with heard/read words, e.g. `platypus` or `green`
  - They also know when words may have more than one meaning or part of speech assigned to it
    - e.g. `duck` (the animal and the act of crouching)
  - The same words don't always convey the same meanings when the order is switched around:
    - e.g. `the green duck dawdled around the cactus` vs. `the duck dawdled around the green cactus`

- *Semantics*: the study of meaning in language

  - There are *natural languages*, *computer languages*, *formal logics*, etc.
  - In all cases, some notion of semantics persists regardless of medium
  - Large field of study: *lexical* semantics, *operational semantics*, *axiomatic* semantics, *denotational* semantics, etc.
  - When we study mathematical linguistics, we will return to many of these definitions!

- Native speakers also are able to use contextual bases for interpreting sentences

  - Ability to judge whether a particular utterance is appropriate or inappropriate in a situation
  - Can discern ideas and exercise discretion based on previous knowledge (time, place, etc.)
  - e.g. "`Can you close the door?`" can be interpreted in two ways:
    - Someone's question as to whether you possess the ability to close a door
    - Someone's request that you actually close a door in proximity
  - Context tells which, but previous usage suggests that the intended meaning would be the latter, as the former would be indicated by a slightly different question, e.g. "`Are you able to close a door?`"

- *Pragmatics*: the study of how context contributes to meaning in a language

  - Deeply related to *semiotics*, the study of meaning-making in a language
  - Similar to semantics, but with a slightly different twist on the meaning-focus:
    - Semantics: how meaning is "coded" in a language, so to speak
    - Pragmatics: how transmission of language depends on structural/linguistic knowledge AS WELL AS context/pre-existing knowledge of the content
  - As briefly mentioned above, pragmatics can help language speakers overcome ambiguity in utterances
  - *Pragmatic competence*: ability to understand another speaker's intended meaning
  - *Pragmatic performance*: demonstrated correct interpretation of a speaker's intended meaning
    - Pragmatic performance errors: misjudging context, e.g. accidentally using the wrong phrase at an inapt time
    - There is a strong tie to sociolinguistics in this aspect of pragmatics ("judgements")

- The written aspect of a language is also important, though less so in today's world

  - Argument can be made that computer-typing is still a similar skill, so it persists
  - Linguists need to study writing systems and how they work to recover lost knowledge from earlier civilizations
  - *Orthography*: set of conventions for written language (e.g. spelling, hyphenation, capitalization, etc.)
  - *Graphemics*: study of writing systems and their basic components (*graphemes*)

- We now understand the various ways to measure linguistic performance, but what about linguistic competence?
- How is the knowledge of language actually stored in the brain?
  - People produce language often, but it is not a tangible object that is measurable
  - Yes, we can "create" sentences, but it's just a transmission of ideas from mind-to-mind
  - That is, **language exists only in the minds of its speakers**
  - i.e. linguistic competence is nothing but language itself!
- A common breakdown of linguistic competence in the mind:
  - *Lexicon*: collection of all words known by speaker, what functions they serve, how they are pronounced, etc.
  - *Mental grammar*: set of rules known implicitly by the speaker about the language
- Linguistics puts a slightly different spin on the terms "grammar" and "rule" when compared to casual conversation:
  - Linguistic grammar = language system
    - Comprises all elements and rules about phonetics, phonology, morphology, syntax, and semantics that make up language
    - Pragmatics and semiotics don't really matter in the linguistic sense of grammar
  - Linguistic rule = statement of pattern that occurs in a language
    - Rules in mental grammar help produce well-formed sentences as well as proper interpretation
    - Here, pragmatics and semiotics play a much larger role, as context matters for rules
- Mental grammar rules may not be the same as written rules
  - Children acquire these rules for their native language unconsciously during their early years
  - As they grow up, these children pay attention to the usage of the language around them
  - That is, short of mental incapacitation, children can acquire the language they are exposed to without being taught
  - Some scholars argue that the plasticity of this ability deteriorates sharply after infancy and childhood
  - Adults are not naturally-equipped for this ability, but by studying linguistics, can learn to learn better
- *Language acquisition*: the study of how language is acquired (by children, by adult learners, etc.)
- Native speakers' mental grammars are unique, i.e. the mappings made by one are not shared by others in general
  - However, they are similar enough that the differences are negligible
  - Very seldom do two people disagree as a consequence—understanding each other as native speakers is standard
- That said, there is certainly *language variation* found from region to region, and even from country to country
- The mental grammar concept forms the backbone of the communication chain discussed earlier
- How do linguists uncover and describe what native speakers know?
  - i.e. need to realize the mental grammars and expose the "hidden knowledge"
  - That is, linguists use linguistic performance to deduce the rules in speakers' linguistic competence
- Standard science: using observable evidence to draw conclusions about unobservable internal structure
- *Descriptive grammar*: collections of generalizations based on observations about a language
  - Linguists use descriptive grammars to determine what mental grammars must consist of
- *Prescriptive grammar*: collections of generalizations about what a language should be like
  - Three notions of "grammar": mental grammar, descriptive grammar, and prescriptive grammar
  - Language is NOT prescriptive grammar!
- Some descriptive observations about English:
  1. The vowel sound in the word `suit` is produced with rounded lips
  2. The sequence of sounds [brt] can form a possible word in English
  3. Adjectives come BEFORE the nouns they describe, e.g. `green shirt`, not `shirt green`

Language Knowledge in Practice

- Writing is secondary to speech, and is not actually necessary for knowledge of a language

    – Evidence: children, who learn to natively speak a language before every learning to write
    – Evidence: pre-historic times, when people passed down information orally before writing systems existed
    – Audio-vocal skills develop at a much younger age than fine motor skills do, anyway

- Speech is typically the immediate manifestation of language for linguistic performance

- However, writing can also be used to express knowledge of a language, as has been the custom for the last few millennia

- As noticed in the topics described in the theory of language knowledge section, the larger focuses is on speech and speaking-related linguistic phenomena when compared to the phenomena that exist in the written medium

- *Writing*: representation of language in a physical medium (different than sound)

    – Speech = immediately physically transmittable form of language
    – Writing = physically preservable form of language
    – Both speech and writing encode thought into communicable forms, but modern technology blurs the distinction

- Writing adds another step to the communication chain, namely the act of writing it down (we omitted this earlier)

- **All units of writing are based on units of speech**

    – e.g. words, *morphemes*, syllables, sounds, etc.
    – i.e. must first process ideas via the speech system and then transfer the ideas into writing

- To measure mental language competence, it is important to come as close to the original source as possible (i.e. speech)

- Why focus on written aspects of language?

    – Not all languages are observable anymore, e.g. Latin (which has been *dead* for a few centuries)
    – In dead languages, the written form **is** the closest source to the original
    – Related to a field of linguistics known as *corpus linguistics*, focusing on real texts

- Speech is a more basic form of language than writing:

    1. Writing must be taught:
        – Spoken language is acquired naturally, before children even begin attending school
        – Spoken languages can even develop spontaneously in societies where a full language may not exist (yet)
        – All writing systems are taught explicitly, usually in some form of schooling
    2. Writing does not exist everywhere:
        – Our societies are highly literate in the West, but this is NOT the case around the world
        – There are ~7100 languages around the world, ~3535 of them are actually *unwritten languages*
        – Some people are illiterate, but are still competent in the spoken form of their native language
        – No society uses a pure written form of language without ever speaking it as well
    3. Neurolinguistic evidence:
        – The production of written language in the brain is overlaid on the same places that spoken language occur
        – Spoken language is fundamental: it uses several distinct areas, but written language uses these and MORE
    4. Writing can be edited:
        – Speech is much more spontaneous than writing, except perhaps in arenas of debate, etc.
        – Suggests the immediacy of speech as a signal for communication (writing is delayed in comparison)

- So why do people believe that writing is a "purer form" when compared to speech?

    1. Writing must be taught:
        – Intimately associated with education and educated speech (erudite aspect of language)
        – In centuries past, literacy was only for the elite/royal classes, i.e. reflects a class divide structure
        – Writing is indirectly associated with the "correct" varieties of language, as it were

– Some writers try to preserve regional dialects by staying faithful in their transcript of speech:
  ○ *The Adventures of Huckleberry Finn*, Mark Twain (1885)
  ○ *Their Eyes Were Watching God*, Zora Neale Hurston (1937)
  ○ *To Kill a Mockingbird*, Harper Lee (1960)

2. Writing can be edited:
   – Written products tend to be more aptly worded and better organized (fewer errors, hesitations, pauses, etc.)
   – Writing is often the result of deliberation, correction, and revision (easier to demonstrate competence)

3. Writing is physically stable:
   – c.f. sound waves travelling through air (*phonons*)
   – Physical media are much more stable than audio/video recordings
   – Spelling does not vary (anymore, as it has been standardized) as much as pronunciation does
     ○ American vs. British vs. Canadian English is an exception
     ○ This is also seen in computer-based writing ("thru", "nite", etc.)

- So the intrinsic characteristics of writing make it seem more polished/permanent, but evidence does suggest that it is NOT a more primary indication of one's linguistic competence

- Regardless, writing still serves as a significant aspect of linguistic study (as it relates to language)

- Descriptive grammar statements vs. prescriptive grammar statements:

  – Descriptive statements:

    1. Some English speakers may end sentences with prepositions.
    2. Some English speakers may split infinitives.
    3. Some English speakers use double negatives for negation or emphasis.

  – Prescriptive statements:

    1. Do not end sentences with prepositions in English.
    2. Do not split infinitives in English.
    3. Do not use double negatives for negation or emphasis in English.

- Prescriptive rules make *value judgements* about the correctness of a phrase

  – Indeed, double negatives are actually REQUIRED in some languages (Spanish, Russian, etc.)
  – e.g. `no, no puedo ir a la playa hoy` (lit. no, no I can go to the beach today)

- However, rules in a native speaker's mental grammar cannot (by definition) be incorrect

  – True even if these rules differ from rules of other speakers' mental grammars
  – True even if these rules differ from GENERAL descriptive grammatical statements about the language

- Descriptive languages, in comparison, allow for different varieties of a language

  – They do not ignore a construction just because a prescriptive grammarian does not like it
  – No limitation on what speakers "can" or "cannot" do, just statements on what they ACTUALLY do
  – e.g. the phrase `nobody ain't done nothin'` is NOT incorrect in English
  – e.g. some dialects of English admit phrases like `the room needs painted`, while others may view this as incorrect and instead say something like `the room needs painting` or even `the room needs to be painted`

- Important to keep in mind: languages change over time!

  – *Northanger Abbey*, Jane Austen (1818): `the clock struck ten while the trunks were carrying down`, as opposed to more current English, where `the clock struck ten while the trunks were being carried down` would be considered more correct
  – i.e. grammaticality can undergo drastic changes in a century, as it is not inherent to a language
    ○ Indeed, older forms of English—Old English and Middle English—had double, triple, and even quadruple negatives (often used for emphasis purposes)
  – It follow that prescriptive grammar CANNOT be the backbone for language
    ○ i.e. prescriptive rules are used only as social identity markers via judgements about linguistic validity

Language Design Features

- How to properly define what a language is, entails, and encompasses?

    – Hard to do, fundamental nature is hard to capture
    – One solution: identify features that something must have to be a language
    – *Design features*: so-called "descriptive characteristics of a language"

- The nine design features of a language (by Charles Hockett, 1960):

    1. Mode of communication:
        – Recall communication chain: messages must be sent and received
        – Mode of communication: means by which messages are transmitted and received
        – Common modes: voices, gestures (hand, arm, facial movements)
        – Later on: the idea of *language modality*

    2. Semanticity:
        – All signals in a communication system must have a meaning/function
        – This is true even when people hear words they don't know (assumption of existence of meaning)
        – i.e. words and sentences must have meanings for us to communicate with them

    3. Pragmatic function:
        – Communication systems must serve a meaningful purpose
        – e.g. helping people stay alive, influencing behavior, discovering more in the world, etc.
        – Are all communicative acts useful? (e.g. gossip)
            ○ However, even gossip serves a useful purpose in society
            ○ Helps people understand social environments and is key in social bonding/relationships
            ○ Similarly for "small-talk" phrases: (A: "Hey, what's up?", B: "Not much, what about you?")

    4. Interchangeability:
        – Ability of individuals to both transmit and receive messages
        – Production/transmission: speaking or gesturing
        – Consumption/reception: listening or watching

    5. Cultural transmission:
        – Some aspects of language can only be acquired through communicative interaction with other system users
        – i.e. though children's ability to acquire language seems innate, they must still learn specific language signals
        – Research shows that children that are not spoken to will not learn language
        – Moreover, children learn the language that other people use to interact with them
            ○ Children of Russian parents will learn Russian if their parents interact with them in Russian
            ○ But they will learn English if their parents interact with them in English instead
            ○ They could even learn Russian-accented English if, e.g. their parents are Russian immigrants to the U.S.
            ○ i.e. **genetic or hereditary background has no influence whatsoever on language acquisition**

    6. Arbitrariness:
        (a) Language:
            – Words represent a connection between a group of sounds/signs
            – This group gives a word a *form* and meaning (the form represents the meaning)
            – *Linguistic sign*: the combination of a form and meaning
            – e.g. "inner core of a peach" is represented by [pɪt], spelled ⟨pit⟩ (must occur in THAT order to give the sound/form that is made when `pit` is said aloud)
            – Will explain the [·] notation later!
                ○ Preview: they are *transcriptions* in the *International Phonetic Alphabet* (IPA)
                ○ IPA: standardized set of symbols used to indicate pronunciations for ALL languages
            – Connection between form and meaning is usually *arbitrary*
            – Arbitrary: meaning is not predictable from form, nor does the meaning dictate the form
            – However, meanings are FIXED based on form, so there is a relationship/connection there, at least
            – *Conventions* describe how a certain group of sounds associate with a particular meaning

- – Opposite concept: *nonarbitrariness*
  - ○ e.g. *iconic* connections, where forms directly represent meanings
- (b) Evidence:
  - – Going back to example: a `pit` can also be called a `seed`, `core`, etc.
  - – Suggests that there is some sense of arbitrariness in the connection between form and meaning
  - – If the connection were non-arbitrary, there would only be a single form to express a meaning (*one-to-one*)
  - – Other evidence: cross-linguistic comparisons
    - ○ Same meaning may have different forms in different languages
    - ○ Same form may have different meanings in different languages
    - ○ i.e. there are no universally-recognized forms for ALL meanings (maybe some?)
  - – More evidence: inventions and new concepts
    - ○ Names of objects and concepts that are new vary widely, even in existing fields
    - ○ The intrinsic properties of these things do NOT dictate their names
    - ○ Rather, the arbitrary preferences of the creator decide the names of these products
- (c) Onomatopoeia:
  - – Also a form of linguistic evidence for arbitrariness
  - – *Onomatopoetic* words: words that are imitative of natural sounds or have meanings associated with similar sounds existent in nature
  - – Examples in English:
    - ○ `splat`: sound of rotten tomato hitting a wall
    - ○ `cock-a-doodle-doo`: sound that rooster crows in the morning
  - – Onomatopoetic forms are LARGELY determined by meaning, but form is NOT an exact copy of noise:
    - ○ That is, roosters do not actually "say" [kɑkədudl̩du]
    - ○ This is merely a convention that English speakers have rationalized
  - – Interesting evidence: different languages can have different onomatopoetic words for the same sounds
  - – This can be seen in the following table:

Sound	English	German	French	Spanish	Hindi	Mandarin	Japanese	Greek
Dog barking	[baʊwaʊ]	[vaʊvaʊ]	[wafwaf]	[waʊwaʊ]	[bʰɔ̃bʰɔ̃]	[waŋwaŋ]	[waɴwaɴ]	[ɣavɣav]
Rooster crowing	[kɑkədudl̩du]	[kikəʁiki]	[kokoʁiko]	[kikiɾiki]	[kukukuku]	[kukuku]	[kokekokkoː]	[kikiriku]
Cat meowing	[miaʊ]	[miaʊ]	[miaʊ]	[miaʊ]	[miaʊ]	[miaʊ]	[niaɯ]	[ɲaʊ]
Cow mooing	[muː]	[mu]	[mø:]	[mu]	[mũː]	[məɹ]	[mo:mo:]	[muː]
Sheep bleating	[baː]	[mɛː]	[bɛː]	[beː]	[mɛ̃ːmɛ̃ː]	[miɛ]	[me:me:]	[beː]
Bird chirping	[twittwit]	[pippip]	[kɥikɥi]	[piopio]	[tʃiːtʃiː]	[tɕitɕi]	[tʃitʃi]	[tsiutsiu]
Bomb exploding	[bum]	[vʁum]	[bum]	[bum]	[bʰɔɖaːm]	[bɔ̃ŋ]	[baɴ]	[bum]
Laughing	[hɑhɑ]	[hatʃa]	[haha]	[xaxa]	[haha]	[xaxa]	[haha]	[xaxa]
Sneezing	[atʃu]	[hatʃi]	[atʃum]	[atʃu]	[atʃũː]	[aʔtʰi]	[hakɯʃoɴ]	[apsu]
Something squishy hitting a hard surface	[splæt]	[platʃ]	[flɔk]	—	—	[pyaʔ]	[guʃaʔ]	[plats]
Clock ticking	[tɪktɑk]	[tɪktak]	[tiktak]	[tiktak]	[tiktik]	[tiʔtaʔ]	[tʃɪktakɯ]	[tiktak]

Figure 1.1: Cross-linguistic examples of onomatopoeia, using IPA symbols.

- (d) Sound symbolism:
  - – Certain sounds occur by not directly imitating but rather by being evocative of a particular meaning
  - – That is, these words ABSTRACTLY suggest physical characteristics based on how they sound
  - – In many languages, words for `small` and that have the concept of "smallness" as part of their meaning often contain the vowel [i]
  - – This is seen in some English words:
    - ○ `teeny`: extra small
    - ○ `petite`: small or thin
    - ○ `wee`: very small
  - – Similarly in other languages:
    - ○ `mikros`: small (Greek)

- ○ `perrito`: little dog (Spanish)
- – For Spanish specifically, this is part of a larger phenomenon:
  - ○ *Diminutive nouns*: words that have been modified to convey a slighter degree of intended root
  - ○ In Spanish, this is done with the suffix `-ito`
  - ○ Spanish also has *double-diminutive* forms: `perro` ⇒ `perrito` ⇒ `perritito` (i.e. `-ito` + `-ito` = `-itito`)
  - ○ Opposite concept: *augmentative nouns* (e.g. `super-` and `grand-` prefixes in English)
- – So [i] is generally associated with "smallness" in languages
- – Japanese sound symbolism (*mimetic words*) is a bit more developed:
  - ○ *Animate phonomime* (擬声語, or *giseigo*)
    - □ Words that mimic sounds made by living things, e.g. a dog's bark
  - ○ *Inanimate phonomime* (擬音語, or *giongo*)
    - □ Words that mimic sounds made by inanimate objects, e.g. wind blowing or rain falling
  - ○ *Phenomime* (擬態語, or *gitaigo*)
    - □ Words that depict states, conditions, or manners of the external world, e.g. "damp" or "stealthily"
  - ○ *Psychomime* (擬情語, or *gijōgo*)
    - □ Words that depict psychological states or bodily feelings
  - ○ Of these, giongo and gitaigo are the most common onomatopoeia
  - ○ In linguistics, these types of onomatopoeia are known as *ideophones*
- (e) Non-arbitrary aspects:
  - – Nonarbitrariness (e.g. iconicity) has a marginal impact on language, at best
  - – However, non-arbitrary aspects certainly play SOME role in shaping a language
  - – Indeed, speakers are aware of their potential effects
  - – Poets sometimes manipulate onomatopoeia and sound symbolism to achieve a specific phonic impression:
    - ○ e.g. *The Princess*, Alfred Tennyson (1847)
    - ○ Utilizes nasal consonants to mimic the noise made by bees (to which he refers often)
    - ○ Example from verse 11, lines 206 and 207 (emphasis ours):
      > The **moan** of doves **in** i**mm**e**m**orial el**ms**
      > And **m**ur**m**uring of i**nn**u**m**erable bees

7. Discreteness:
   - – Property of language that lets us combine together discrete units to create larger communicative units/blocks
   - – e.g. `he is fast` is composed of the independent words `he`, `is`, and `fast`, which in turn are composed of the individual sounds [h], [i], [ɪ], [z], [f], [æ], [s], and [t]
   - – All languages have a limited number of sounds ($10 \lessapprox n \lessapprox 100$ usually, where n is the number of sounds)
   - – e.g. for English, $n \approx 50$
   - – These sounds in and of themselves are universally considered meaningless
   - – However, a small number of these sounds can be combined to create a large lexicon of meaningful words
   - – Can combine or reorder sounds to get completely different words:
     - ○ [p] + [u] + [l] = pool; [k] + [u] + [l] = cool
     - ○ [kul] = cool; [klu] = clue; [luk] = Luke (proper name)
   - – Meaningful elements: words; meaningless elements: sounds
   - – An infinite number of meanings can be created using these few meaningless sounds (if no restriction on length)
   - – Limiting the length ⇒ limiting number of possible meanings ⇒ a less useful language for communication!

8. Displacement:
   - – Ability of a language to communicate about actions/ideas that are not present in the space/time where the speakers are currently communicating
   - – e.g. the ability to talk about the color red without actually seeing it
   - – Can also talk about things that do not or cannot exist (unicorns, fictional characters, etc.)

9. Productivity:
   - – Capacity of language for novel messages to be built up out of discrete messages
   - – Productivity vs. discreteness:
     - ○ Discreteness: communication system must have recombinable units
     - ○ Productivity: no fixed set of ways in which units can recombine

○ Productivity is not as limited as discreteness, as the latter can enforce a fixed set of *recombination rules*
- Some communication systems work in limitation, but language is productive, so there is no fixed set
- Productivity grants people ability to produce/understand an infinite number of previously-unheard sentences
- e.g. probably have never seen this sentence before: `funky potato farmers dissolve glass`
    ○ May not know why potato farmers are funky or how they dissolve glass
    ○ But have an idea of what sentence encompasses and an image of what is happening
- Construction: putting together sounds, morphemes, and words in regular, systematic, rule-governed ways
- Rules at ALL levels of linguistic structure are productive
    ○ i.e. linguistic rules are all levels allow creation of new forms and explain how to use them
    ○ Rules of language are NOT limitations—rather, they open up ability to broadly communicate ideas

- Some other design features that Hockett expanded upon in 1966 (more on these later):

    10. Broadcast transmission/directional reception
        - When humans speak, sounds are transmitted in all directions, but listeners only perceive from one direction
    11. Transitoriness
        - Language quality is temporary, i.e. language sounds only exist for a brief period of time
    12. Total feedback
        - Speakers of a language can hear their own speech and control/modify what they are saying in real time
    13. Specialization
        - Purpose of signals is for communication, not another biological function
    14. Prevarication
        - Humans have the ability to lie or deceive with language
    15. Reflexiveness
        - Humans have the ability to use language to talk about language
    16. Learnability
        - Language is teachable to and learnable by humans

- All languages (human, computer, etc.) that we know exhibit these nine design features

    - **Any communication system that does not exhibit all of these design features is not a language**
    - Moreover, only HUMAN languages appear to exhibit all of these characteristics (language is uniquely human)

- However, if a communication system DOES exhibit all nine features, is it necessarily a language?

    - *Formal language*: a language consisting of a set of strings of symbols, together with a set of governing rules
        ○ e.g. formal languages used to write mathematical proofs, computer languages used for programming systems
        ○ Formal languages fundamentally differ from languages like English, Spanish, Mandarin Chinese, etc.
        ○ No child can acquire OCaml as a native language!
        ○ But what about constructing artificial languages for fun/media (Dothraki, Klingon, Esperanto, etc.)?
        ○ Is there a distinction to be made between English/Spanish/Mandarin Chinese and Dothraki/Klingon/Esperanto?
        ○ Questions are open to debate/research, but some argue "yes": natural vs. constructed languages
    - *Natural language*: a language that has evolved naturally in a speech-based community
    - *Constructed language* (or the portmanteau, *conlang*): a language specifically invented that may or may not imitate all of the properties of a naturally-occurring language (sometimes called *natlang* by analogy)
    - Some constructed languages have the potential to become natural languages, if learned by native speakers
    - This happened with Modern Hebrew: reconstructed from Ancient Hebrew, grew popular in the early 1900s
        ○ This type of language is sometimes called a *reconstructed language*
        ○ However, Hebrew has been the only instance of a successful reconstruction (other attempts: Latin, etc.)
    - Conlangs vs. formal languages: formal languages cannot be acquired naturally
    - We will mostly restrict our study to only natural languages henceforth (i.e. "language" = "natural language")

Language Modality

- Insufficient to purely use the "mental grammar" model to describe mode of communication

- Another approach: introducing the idea of how a language is produced and perceived

- *Modality*: mode of communication in a language

- Most well-studied natural languages are *auditory-vocal languages*

    - Sometimes called *auditory-oral languages* or *spoken languages*
    - These languages are perceived via hearing and produced via speech
    - e.g. English, Russian, Portuguese, Navajo, Korean, Swahili, etc.

- By comparison, some human languages are *visual-gestural languages*

    - Sometimes called *signed languages* (SLs)
    - Perceived visually and produced via hand/arm movements, facial expressions, head movements, etc.
    - e.g. American Sign Language (ASL), etc.

- Auditory-vocal and visual-gestural languages make up the vast majority of human languages

    - However, there are a few other modes of communication as well
    - e.g. deaf-blind people may use a tactile-gestural modality (Helen Keller, etc.)
    - Tactile-gestural: perceived by feeling, produced similar to visual-gestural languages
    - Alteration in modality here does NOT mean that this is a new type of language
    - e.g. SLs can be used as tactile-gestural languages (if motions are felt)
    - That is, SLs have a *primary modality* (visual-gestural) and *secondary modality* (tactile-gestural)

- SL vs. manual codes:

    - Myth: SLs are derived from spoken languages, i.e. they are just "coded"
    - Codes and languages are very different types of systems, however
    - *Code*: artificially constructed system for representing a natural language
    - Codes have no structure of their own, but borrow one from the represented natural language
    - SLs evolve naturally, independently of spoken language counterparts, if any
    - Evidence: British Sign Language (BSL) and ASL are completely unrelated
        - People fluent in BSL cannot understand ASL necessarily, and vice versa
        - OTOH, people fluent in American English can understand people fluent in British English quite well
    - However, manual codes for spoken languages DO exist
        - e.g. Signed Exact English II (SEE II) vs. ASL
        - Codes may represent letters, morphemes, words, etc.
        - Codes may even follow the grammar!
        - However, information is usually transmitted faster in SLs when compared to manual codes
        - We will not focus very much on manual codes

- SL vs. pantomime:

    - *Pantomime*: a form of signed communication where signers using hands to draw pictures in air or act out thoughts
    - SLs are NOT pantomimes, though
    - SLs have an internal structure governed by phonological, morphological, and syntactic rules
    - Indeed, forms in SLs are also arbitrary in terms of relationship to meaning
        - A particular gesture means `possible` in ASL but `weigh` in Finnish Sigh Language (FSL)
        - Arbitrariness is present even in SLs
        - Just like how words may be pronounced differently regionally, signs may be signed differently regionally
    - Interesting: signs may start out as iconic in a SL, but eventually change over time (arbitrariness)

- Universality of SLs:
  - SLs are not universal at all! See above examples (ASL vs. BSL vs. FSL, etc.)
  - Pantomime has a larger degree of universality, as it is iconic in nature
  - Indeed, there are ∼150 documented SLs in the world!
- Representing signs on paper:
  - i.e. a two-dimensional format representation of signs (a meta-representation of language)
  - Convention: spelling out signs using an English word in capital letters to represent the sign for that word
  - e.g. the ASL sign for dog may be written on paper as DOG
  - Can view three-dimensional versions online: http://linguistics.osu.edu/research/pubs/lang-files/links
- Why study different modalities?
  - **Modalities allow us to understand the core linguistic principles from a variety of perspectives**
    - Indeed, allows us to better understand language itself!
    - Can observe what sorts of effects modalities have (or do not have) on language
  - Articulation and perception are completely different in different modalities
  - However, most of our study will concentrate on spoken languages henceforth
- We will principally focus on the auditory-vocal modality henceforth
  - Applications to visual-gestural and other modes of communication will be presented sparingly
  - That is NOT to say that speech is the "best" form of language (this is a value judgement)
  - Just easier to focus on this due to available resources (research papers in peer-reviewed journals, etc.)

Further Reading

1. Crystal, David. 2010. *The Cambridge Encyclopedia of Language*. 3ed. Cambridge: Cambridge University Press.

2. Groce, Nora Ellen. 1985. *Everyone Here Spoke Sign Language: Hereditary Deafness on Martha's Vineyard*. Cambridge, MA: Harvard University Press.

3. Hall, Christopher J. 2006. *An Introduction to Language and Linguistics: Breaking the Language Spell*. New York: Continuum.

4. Lewis, M. Paul; Gary F. Simons; and Charles D. Fennig. 2015. *Ethnologue: Languages of the World*. 18ed. Dallas, TX: SIL International. Online version: http://www.ethnologue.com

5. Valli, Clayton; Ceil Lucas; Kristin J. Mulrooney; and Miako Villanueva. 2011. *Linguistics of American Sign Language: An Introduction*. 5ed. Washington, DC: Gallaudet University Press.