SCIENTIFIC METHODS AND EMPIRICISM
Empiricism is a cornerstone of the **scientific method**, but it is **not** the only way of discovering truth (e.g. does John love Mary, are sunsets beautiful?). It is, however, the approach that we will use to analyze linguistic phenomena in this class.
Definition

The **scientific method** is defined by:

1. Empirical, systematic observation
2. Publicly available knowledge
3. Dealing with problems that are solvable and potentially falsifiable
1a. Empirical observation

- **Empirical** < Gk. empeirikós: experienced/tested
- **Rational(ism)** < Lat. ratio: reason

- Theories should be based on our observations of the world rather than on intuition, faith, reasoning, or appeals to authority
Example of non-empirical

- People refused to look through Galileo’s telescope to see Jupiter as a new planet (refused empirical evidence). Francesco Sizi refuted Galileo by using “reasoning” his reasoning was:

  “There are seven windows in the head, two nostrils, two ears, two eyes and a mouth; so in the heavens there are two favorable stars, two unpropitious, two luminaries, and Mercury alone undecided and indifferent. From which and many other similar phenomena of nature such as the seven metals, etc., which it were tedious to enumerate, we gather that the number of planets is necessarily seven.... Besides, the Jews and other ancient nations, as well as modern Europeans, have adopted the division of the week into seven days, and have named them from the seven planets: now if we increase the number of planets, this whole system falls to the ground.... Moreover, the satellites are invisible to the naked eye and therefore can have no influence on the earth and therefore would be useless and therefore do not exist.”
Examples of modern day Sizis

- Too many linguistic arguments appeal to Chomsky’s prestige and authority rather than data.
- Chomsky’s refusal to accept what is “observed” in a corpus as relevant
- Conference presentation refuting the Spanish philologist Menendez Pidal
1b. Systematic observation

- Observation alone is not enough, it must be systematic. Observing everything you do one day doesn’t constitute a systematic observation. You need to observe things that are relevant to the theory and are structured so that they can either support or refute the theory.
2. Scientific knowledge is publicly available

Examples:
• Cold fusion
• South Korean cloning experiments
• String theory
• Chomsky and introspection (example of *perform + mass noun: "perform magic / labor")
• If knowledge is not available publicly it can never be scrutinized, examined, critiqued, or refuted like public knowledge can. Nor can it be replicated.

Importance of peer reviewed publication-makes it public, subject to scrutiny. Just because it’s been peer reviewed doesn’t mean it’s true. It a minimal standard.

You should be wary of anything that hasn’t been studied and published: diet pills, megavitamins that cure schizophrenia, depression. Acupuncture was accepted as legitimate only after it was tested.
3. Testable problems / potentially falsifiable

- Falsifiability is good, as strange as that may seem. It’s OK for a theory to be proven wrong. We get closer to truth.
- Science deals with theories that can be tested. The test must be based on spatiotemporal evidence-observable (not appeals to authority, reasoning)

Scientific theory must specify what outcome would support and what would disprove it.

Not:
- What is the meaning of life?
- Are humans inherently good or bad?
- Was Monet the greatest 19th century painter?

- Example 1: Dr. Rushmore and the 1793 yellow fever epidemic in Philadelphia
- Example 2: /r/ flaps when followed by a [-continuant] consonant (butter, pedal, barter). What is [l]? (faculty, altar)
Are non-scientific enterprises good? Yes; poetry, art appreciation, musical preferences
Scientific Method

• Observe something that gets you thinking about why it is so, or what makes it that way.
• Form a hypothesis to explain it. The hypothesis must state what would verify or disprove it.
• With an experiment or other observations test the hypothesis.
• Draw conclusions based on results, which may entail rejecting or modifying hypothesis.
• Make the results public by presenting and publishing.
How could this be carried out for the following (or in fact could it?):

• Passives occur more in English than in Spanish.
• Floridians speak better than people from Mississippi.
• The vocabulary in Shakespeare in larger than that of the King James Bible.
• Poetry has "prettier" language than academic textbooks.
• Money makes people happy.
• Humans have an innate sense that murder is bad.
• Monet was the most prolific painter of the 19th century.
• All languages descend from the language spoken by Adam.
• English is harder for native English children to learn than Chinese is for native Chinese children.
• The words *have*, *be*, and *go* are used about as often in British and American English.
• The word *was* in Shakespeare’s time was pronounced with the same vowel as the word *loss*. 
EARLY EMPIRICAL LINGUISTICS

• Kaeding (1897) had 5,000 people compiling a corpus of 11 million German words (and calculating their frequency, distribution of letters).
• Eaton (1940) compared the frequency of words in several European languages.
• In the U.S. linguistics was tied to anthropology. Focus was on documenting and writing grammars for American Indian languages.
• Behaviorism (Skinner/Bloomfield) focused on stimulus response (Skinner boxes). Mental processes are unobservable so they are not studied.
EMPIRICISM DURING THE PAST 50 YEARS

1. Syntactic structures (1957) by Noam Chomsky started talking about mental processes of language.
2. Argued that language acquisition can't be response to stimuli. There is no reward/punishment for “correctness”. Poverty of the stimulus argument shows something beyond what is heard must account for language abilities.
3. Chomsky led move from empiricism to rationalism:
   - Why worry about getting 10,000,000,000 word corpus? Just ask native speaker.
   - Empirical data often non overly-insightful (Dayton, OH vs New York City)
   - (vs behaviorism): Degenerate data: motherese, competence/performance (i.e. drunk people)
   - Analogy of human body: just study one in detail
   - Can a corpus really tell you why "he shine Dad boots" is bad?

So empirical linguistics really "on the ropes" late 1950s-late 1970s or into 1980s
EMPIRICISM DURING THE PAST 50 YEARS (2)

4. Resurgence of empirical studies since 1980s or so. Why?
   - Computers could finally handle enough data to make real contribution
   - Problems with introspective data:
     - It can’t be applied to acquisition (Is cow is noun or verb?)
     - It doesn’t admit variation (going, goin’; gave to John the book/gave John the book)
     - How can you be sure of or verify someone’s introspection. What if 2 people’s introspections differ? Introspections are subjective not objective.
     - Introspections don’t allow frequency to be measured.
     - Introspections allow people to make up data to fit their theory, whether consciously or subconsciously