

AMBIGUITY

From "The Cambridge Encyclopedia of the Language Sciences"

 CAMBRIDGE UNIVERSITY PRESS © Cambridge University Press 2011

Ambiguity refers to the potential of a linguistic expression to have more than one meaning. Although many expressions (words, phrases, and even sentences) are ambiguous in isolation, few remain so when used in a particular context. In fact, people typically resolve all ambiguities without even detecting the potential for other interpretations. Ambiguity does not imply **vagueness**; rather, ambiguity gives rise to competing interpretations, each of which can be perfectly concrete. Although ambiguity is pervasive and unavoidable in natural languages, artificial languages developed for mathematics, logic, and computer programming strive to eliminate it from their expressions.

Ambiguity can be lexical, structural, referential, scopal, or phonetic. The examples of these phenomena that follow include well-known classics in English.

Lexical ambiguity refers to the fact that some words, as written or spoken, can be used in different parts of speech (see **word classes**) and/or with different meanings. For example, **duck** can be used as a noun or a verb and, as a noun, can refer to a live animal or its meat. Structural ambiguity arises when different syntactic parses give rise to different interpretations. For example, in addition to being lexically ambiguous, **They saw her duck** is also structurally ambiguous:

1. They saw [_{NP} her duck] (the bird or its meat belongs to her)
2. They saw [_{NP} her] [_{VP} duck] (the ducking is an action she carries out)

A common source of structural ambiguity involves the attachment site for prepositional phrases, which can be at the level of the nearest noun phrase (NP) or the clause. In the sentence **Danny saw the man with the telescope**, either Danny used the telescope to help him see the man (3) or the man whom Danny saw had a telescope (4).

3. Danny [_{VP} saw [_{NP} the man] [_{PP} with the telescope]]
4. Danny [_{VP} saw [_{NP} the man [_{PP} with the telescope]]]

Referential ambiguity occurs when it is not clear which entity in a context is being referred to by the given linguistic expression. Although **deictics** (see **deixis**), such as pronouns, are typical sources of referential ambiguity, full noun phrases and proper nouns can also give rise to it.

5. (at a boy's soccer game) "**He** kicked **him!**" "Who kicked who?"
6. (at a boat race) "**That boat** seems to be pulling ahead." "Which one?"
7. (in a university corridor) "**I'm off to meet with Dr. Sullivan.**" "**Chemistry or math?**" (There are two Dr. Sullivans in different departments.)

Scopal ambiguity occurs when a sentence contains more than one quantified NP and the interpretation

depends on the relative scopes of the quantifiers. For example, **Some children saw both plays** can mean that a) there exist some children such that each of them saw both plays or b) both plays were such that each, individually, was seen by some children but not necessarily the same children. Phonetic ambiguity arises when a given sound pattern can convey different words, for example, **two ~ too ~ to; new deal ~ nude eel**.

Although people typically do not notice the ambiguities that they effortlessly resolve through context, they are certainly aware of the potential for ambiguity in language. In fact, such awareness is a precondition for getting the joke in Abbott and Costello's "Who's on First?" skit or in headlines like "Iraqi Head Seeks Arms."

Whereas ambiguity does not frequently hinder effective communication among people, it is among the biggest hurdles for the machine processing of language. This is not surprising if one considers how much reasoning is required to resolve ambiguity and how much knowledge of language, the context, and the world must underpin such reasoning. As an example of the large scale of the task, consider the short sentence **The coach lost a set**, which you probably interpreted to mean "the person who is the trainer of some athletic team experienced the loss of a part of a match in an athletic competition" (whether the coach was playing or the team was playing is yet another ambiguity). Other interpretations are also valid, given specific contexts. For example, the person who is the trainer of some team might have lost a set of objects (keys, golf clubs) or a railroad car might have lost a set of objects (door handles, ball bearings). If this sentence were used as input to an English-Russian machine translation system that relied on a standard English-Russian dictionary, that system would have to select from among 15 senses of **coach**, 11 senses of **lose**, and 91 senses of **set** – a grand total of 15,015 combinations, if no further knowledge were brought to bear. Of course, all machine translation systems incorporate some heuristic knowledge, and lexicons developed for natural language processing typically do not permit the amount of sense splitting found in dictionaries for people. On the other hand, it is common for sentences to contain upward of 20 words, in which case there is still the threat of combinatorial explosion.

Marjorie J. McShane