

Set Theory for Language Learners

Josef MESSERKLINGER

Introduction

Grammar knowledge is certainly an important feature of language, but semantic knowledge is often of greater concern to us when we try to learn a foreign language. Most of us carry a dictionary around so we can look up unfamiliar words. Once we get the translation(s), we feel that we can work out what message a group of words is trying to convey. For example, in the sentence:

The dishes are on the table.

...once we know what *dishes*, *table* and *on* mean, ignoring the word *the* and assuming we understand *are*, we form an image of objects labeled *dishes* and *table* and the notion that the word *on* describes the relative position of the two objects. From there, we can assume what is where, unfortunately without thinking much more about the grammar of the words or paying much attention to their meaning until we have to look them up again.

Nonetheless, most of us continue to use this strategy because it works most of the time since the relationship between the concepts is clear—the dishes are on the table. Likewise, in this exchange ...

Where are the dishes ?

On the table.

... the question can be easily parsed semantically; the word *where* (and of course its position in the utterance) alerts us that a question about location follows; the symbol at the end (or the intonation) of the string of words confirms this. The sole noun in the question (linked to the adverb with a verb and marked by the article preceding it) makes clear what is being asked about. In the reply, the location word (the preposition) and noun make the location equally clear. The verb, if we think about it, is associated with a present condition, so the speaker/writer really wants to know, *where are the dishes now ?*

But if we didn't know one or more of the words, we would have to guess at what is meant. For example, if we didn't know the word *table*, we might think of places where we normally find dishes besides tables—cupboards, sinks, counters, in some cases the refrigerator, and perhaps even shipping boxes just to name a few other obvious places they could be. Context might give us a clue, for example, if the exchange took place immediately before a meal or right after a meal or when someone has moved into a new house. But the more that is unknown, fewer words and less of the context, the harder it will be to guess, and grammar knowledge will only help us so much. Furthermore, when the common translation and the word differ slightly in concept, confusion can result. Does the word “*dishes*” mean plates and saucers etc. (食器) or food items (料理) as in *the dishes I've prepared for supper* ?

Semantic Parsing

Although the above is a very simple example, this seems to be how many of us try to understand written texts; rather than parse the sentence using grammar rules, we tend to write the translation above each word on the page and then interpret using a bit of common sense. In a more challenging utterance, though, we may have trouble deciding which concept to attach to the word we want to translate. For example, the word *green* refers to a color and may either be the color itself—*Green is my favorite color*—or a word that describes what something else looks like especially when it comes before a noun—*green tomatoes*.¹ We know that leaves in spring and summer can be green as can grass, spinach, many frogs, and American dollar bills, etc. But it can also refer to a place—*The ball landed on the green*. On the other hand, the word *idea* is a noun that refers to something in the mind and like other things in the mind—memories, for example—are usually not thought of as being green, unless meant figuratively as in immature, an association in many languages² with unripe fruit which is usually green. Furthermore, ideas don't normally sleep—things like people, babies, cats, etc. do—but ideas can imply, lead, or idiomatically, take off—unless again sleep is meant figuratively as in lay dormant. Furthermore, although we can stretch our imaginations and picture sleep being done furiously, sleeping fitfully or restlessly is how we might usually express the idea. So, in the well-known example . . .

Green ideas sleep furiously.

...although we can parse the grammar of the sentence quite easily—green is an adjective that modifies the noun ideas which acts as the sentence's subject, etc.—parsing the meaning is quite another problem. And yet, we try to make sense of it.

¹ In French, where the adjective comes after the noun, *tomates vertes*.

² In Japanese, unripe fruit is 青い果物, “blue fruit”, according to Google translate. But 青い can also be translated as inexperienced.

Does it mean that money-making ideas which lay dormant demand to be considered? Or environmental concepts, if ignored, will wreck havoc on mankind?³ We cannot be sure. In this instance, it is the meaning of the words not the grammar that causes difficulty. Here, the typical meaning and usage of each word is limited to a semantic field to the exclusion of the one that follows, making the meaning difficult to understand. Some things just don't naturally go together.

Semantic Limits

This is not to say that grammar is unimportant. We must know not only the meaning of the words we read, but also how they are related to each other and how each functions in the sentence. For example, in this sentence ...

Fruit flies like a banana.

... we must realize that *flies* is a noun and *like* the verb, otherwise we come up with the unlikely image of apples, oranges and other kinds of fruit having the same success with flight that any banana has.

Nonetheless, even in this example, although readers may at first be confused as to whether the subject of the sentence is fruit or flies, given the context, the meaning can eventually be made out.

There is a kind of logic to an utterance that encompasses more than just grammar⁴, but includes our schema, or our conceptual understanding of the world. A problem occurs, however, when we go from concept to word or from word to concept via the apparent short cut of translation. For example, the word *naïve* can be taken to mean *simple* (単純⁵) or to mean the same as the Japanese word ナイーブ, *without frills* or *straightforward*, when more commonly it is used in English to mean "lacking knowledge, experience of life, or good judgment, and willing to believe that people always tell you the truth."⁶ So, in a sentence such as ...

*I'm too busy acting like I am not naïve/ I've seen it all, I was here first.*⁷

... while *simple* might fit, the meaning of the loanword in Japanese makes the quote hard to interpret. *Inexperienced and blindly trusting* makes much more sense. In the above utterance, there is a connection between someone acting like they are not

³ http://en.wikipedia.org/wiki/Colorless_green_ideas_sleep_furiously#Attempts_at_meaningful_interpretations (retrieved October 6, 2014)

⁴ Certainly there are examples where grammar does matter quite a lot: compare *I see what I eat to I eat what I see*.

⁵ The first translation given at <http://dictionary.goo.ne.jp/srch/ej/naive/m0u/> (retrieved September 10, 2014)

⁶ http://www.oxfordlearnersdictionaries.com/us/definition/american_english/naive (retrieved September 10, 2014)

⁷ Cobain, Kurt. "Very Ape." In Utero. CD. DGC. 1993.

naïve—implying that they really are in fact naïve—and claiming to have seen it all. By understanding the situation we will be better able to guess the intended meaning.

So, rather than settling for the first translation in our dictionaries, if we are to understand a text fluently, we need to understand how words in the language we are studying are used together and often appear alongside each other and how the vocabulary of a language fits together. Taking a moment to think about the connection between words may also help us to remember them better.

The Set of All Words

To avoid having to define what a word is, let's assume that the universal set of words in a language is called its vocabulary and a word is any member of that set.⁸ We can group this universal set of words into subsets depending on the function we use to define the set. For example, words can be put into sets according to their function as a part of speech—in English: nouns, verbs, pronouns, etc. So, when asked to list, say, nouns—words that name a person, place, thing, etc.—we might give a list that includes {cat, moon, happiness}. A word like *jumped* does not belong. We can also make sets based on the grammatical function of words, how they are used in an utterance: actors/subjects and verbs/actions, etc. So, in this utterance ...

The moon is beautiful tonight.

... we understand that “moon” is a member of the set of nouns and also a member of the set of subjects, but in this utterance ...

The cow jumped over the moon.

... *moon* is a member of the set of objects (in this case, of the preposition *over*) while *jumped* belongs to the set of verbs.

Added to these subsets of the vocabulary, words can also be members of sets using their meaning as a function. So the image created by this utterance ...

The moon jumped over the cow.

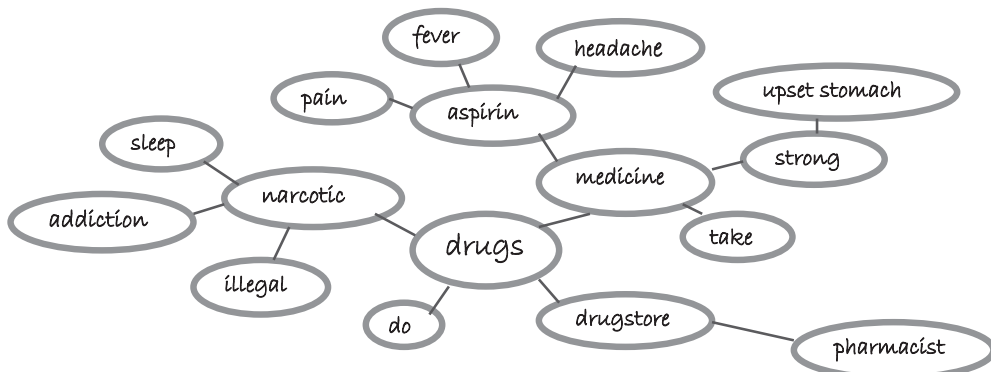
... seems somehow wrong if not merely highly unusual perhaps because it is odd to think of the moon as jumping, although in this utterance ...

⁸ Wikipedia defines a person's vocabulary as “...the set of words within a language that are familiar to that person.” We might also say then that the set of all words in a language is its vocabulary.
<https://en.wikipedia.org/wiki/Vocabulary> (retrieved August 12, 2014)

The moon sailed through the evening skies.

...the verb is perfectly natural because there is a similarity between the way the moon appears to move through cloudy skies and the way a sailboat moves across a calm sea.⁹

This knowledge of which words constitute a set is central to using the language and seems to be how we organize words in our memory. Semantic networks, which are “used when one has knowledge that is best understood as a set of concepts that are related to one another,”¹⁰ is a way to diagram our semantic knowledge. For example, a possible diagram of concepts related to the word *drugs* might look like this:



The diagram above is arranged simply by the associations between the concepts and is not intended to reflect taxonomical hierarchies; it is simply an example of one language user’s intuitive sense of which concepts go together and leaves out those that do not belong. The words make a kind of semantic set. For example, while a word like *movie* may eventually be connected to drugs through its association with common depictions of hospitals or drug abuse in film, a word like *rain* will have a very distant connection to the word *drug* if any at all. Those words belong more commonly to a different semantic set.¹¹ On the other hand, the verb *do* is connected to drugs from the common idiom, *to do drugs*, ie. “to use illegal drugs habitually.”¹²

Sets of Words

The words in the semantic network above are words that go together. Besides their

⁹ the basic meaning of sail is to “travel through water, air; glide” (<http://thesaurus.com/browse/sail>; retrieved August 11 2014.)

¹⁰ https://en.wikipedia.org/wiki/Semantic_network (retrieved August 14, 2014)

¹¹ Besides grouping words into semantic sets, we create pragmatic sets, for example adjacency pairs such as:

A: How are you ?

B: I’m fine thank you.

But pragmatics is beyond the scope of this paper.

¹² <http://idioms.thefreedictionary.com/do+drugs> (retrieved October 21)

conceptual connections—for example, aspirin is a common medicine used for headache and we speak of a *drug addiction* and can say *strong medicine*, but *powerful drugs*. We also say *take medicine*, but, as noted, *do drugs*.¹³ However, these are words that go together in English, and caution must be taken when translating from a semantic net in one language to that in another. For example, when trying to translate *do drugs*, we must decide what word to use for *do* and might come up with:

薬を行う

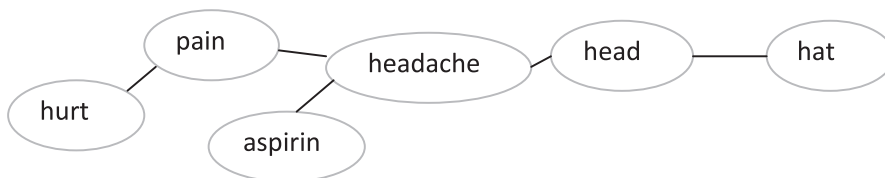
Not terribly helpful. For one, we have the wrong sense in 薬 and miss the distinction between *drug*, a general term for either “a substance used as a medicine” or “an illegal and often harmful substance that people take for pleasure.”¹⁴ *Medicine* is a more specific term that means “a substance that is used in treating disease or relieving pain.” On the other hand, 行う, like *do*, is a very ambiguous word that does not help to make the meaning any clearer and begs the question, how? How are drugs being done?

And in the case of idioms, the relationship between the two words is, as Wikipedia’s explanation suggests, based simply on common use and nothing more. They must be learned as a set, and one way to help learners visualizing these sets is by having them make semantic nets like the one above.

Using Semantic Sets for Pre-reading

Students were instructed how to make semantic nets and then asked to make one around the word pain. The purpose of the activity was to prepare them for an article about improvements in pain treatment. On their first attempt, Set 1, students were told to simply make a network of words from associations with the word pain. The following are typical of each set:

Set 1



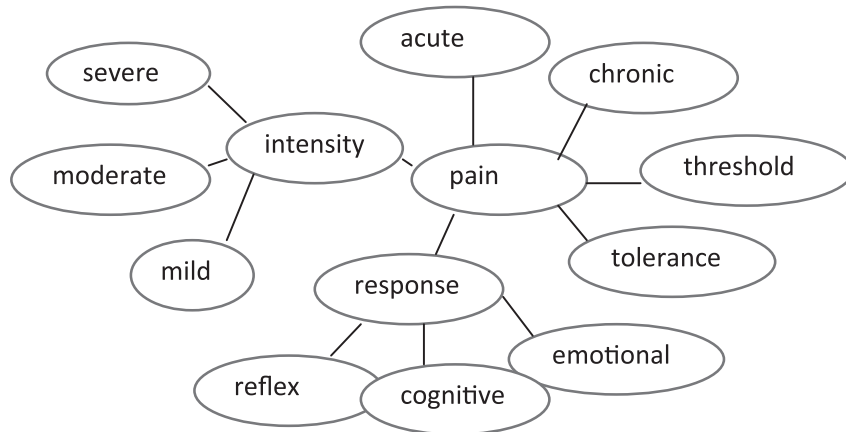
Students were invited to compare their nets and found that although they had some words in common, they each had unique associations which led them to a wide variety of places such as *Nara*, *ghost*, and *zoo*.

¹³ Meaning “use recreational drugs” <http://www.thefreedictionary.com/do+drugs> (retrieved September 13, 2014)

¹⁴ From <http://www.merriam-webster.com/dictionary/drug> (retrieved September 13, 2014)

Next, they were told to limit their nets to the vocabulary from the articles about pain that had been studied during the semester, Set 2. Students were invited to compare again, and this time most came up with semantic nets that looked like this:

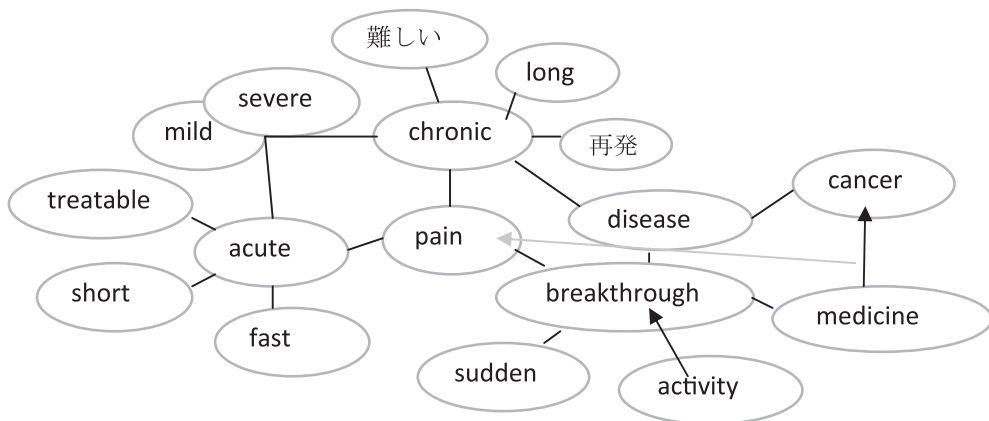
Set 2



The activity allowed students to think again about what they had read and acted as a review of vocabulary. Perhaps not surprisingly, the words many of them remembered were the words studied most intensively in class.

Finally, students were encouraged to use semantic nets as a way of taking notes from a reading passage.¹⁵ With a few prompts, all students were able to make this diagram:

Set 2



¹⁵ <http://learningenglish.voanews.com/content/how-pain-treatment-has-improved-in-recent-years-92695089/114892.html>
(First retrieved March 6 2014)

From this semantic net, students were able to give a concise summary of the reading passage and remember the content and new vocabulary with greater ease. When quizzed the following week on the different kinds of pain, most students were able to answer using the key words from the semantic net.

Conclusion

Semantic nets are the dictionary we carry around with us in our head. But rather than arranged alphabetically, our semantic nets are arranged by the associations between words and either the images the word brings up or the associations themselves. The associations can be based on our understanding of the concepts represented by the words as well as the functions the words perform in utterances. Helping students become aware of this dictionary can help them learn the language they are studying and reinforce vocabulary learning. Making semantic nets can also be used as a way of taking notes, a less restrictive alternative to outlining.