I am honored to participate in inaugurating this new journal, Korea Journal of Chinese Language and Literature. The two languages, Korean and Chinese, indeed the two cultures, share a heritage that reaches back several thousand years. An international journal focused on this common heritage will surely contribute significantly to our understanding of the evolution of language and culture in East Asia.

* I thank the editor of the Korea Journal of Chinese Language and Literature, Professor Ik-sang Eom, for his kind invitation to contribute, and Ms. Yaching Tsai for numerous conversations on the issues discussed in this essay as well as for the Italian examples used herein.
0. Vagueness, Generality, and Ambiguity

The present essay follows the lead of a paper by Y. R. Chao 赵元任 published in 1959 in a volume dedicated to the Swedish sinologist Bernhard Karlgren 高本汉. Although the paper is titled Ambiguity in Chinese and analyzes numerous Chinese examples, it actually provides a framework for viewing ambiguity in general, as opposed to two related linguistic phenomena, i.e., generality and vagueness.

Vagueness is a consequence of using language, which is a discrete code, to represent a world, which is largely continuous. Chao's example of the word brown illustrates this well, since people who speak the same language will vary on how much of the color spectrum to include under this word. Generality is more language sensitive. Chao's example is the Chinese word 母, which is more general than either man or woman in English, though the former is often used in a more general sense; Darwin's classic, The Descent of Man, certainly applies to the evolution of woman as well. Similarly, English friend is more general than Italian amico and amica, the former being male and the latter being female. Whereas I had dinner with a friend in English does not require you to reveal the gender of the friend, this is not an option for Italian.

A highly asymmetric case between English and Chinese is the word cousin, which corresponds to eight terms in Chinese, according to whether the person is male or female, older or younger than yourself, and related on the paternal or maternal side. These terms are 堂兄, 堂弟, 表兄, 表弟, 堂姐, 堂妹, 表姐, and 表妹. No doubt a rich lexicon of kinship terms arose in a culture where big extended families constituted an important social unit for many centuries. Even today, kinship terms are used to address nonkins in Chinese culture to express good will, such as 阿伯, 阿姨, X哥, X姐, etc., where X stands for the person's given name.

Ambiguity refers to the language-specific phenomenon in which a linguistic form refers to two or more meanings, or interpretations, or readings. The meanings may or may not be semantically related to each other. It is a feature of all languages, though there are similarities as well as differences in how it is manifested across languages. The phenomenon is much more prevalent than most of us realize since most of the time it is resolved by context without our ever being conscious of it having occurred. The context in which the form occurs typically disambiguates it, whether the context be linguistic or other.

However, sometimes the ambiguity is deliberate, as is famously illustrated by the oracle at Delphi in ancient Greece. The question put by Croesus, king of Lydia in ancient times, was: 'Should I make war on the Persians?'. The answer was: 'If you make war on the Persians, you will destroy a great realm'. Croesus attacked and lost the war. He should have noted that the oracle was ambiguous with respect to whose 'great realm' would be destroyed; it turned out to be his own! Deliberate ambiguity has a wide range of uses, from enriching imagery in literature, to jokes built on puns, to misleading the consumer in advertising.

Not all ambiguities have such momentous consequences as the Delphic oracle of course. I remember a long and confusing conversation I once had with a Chinese student regarding some phonetic experiments she was doing. I was urging her to find the reasons 原因 for her measurements while she thought I wanted her to segment the vowels 元音 in her acoustic data. This is a case of homophony, since both 原因 and 元音 are pronounced yuányīn.

Another interesting instance was a call I received in Hong Kong from a friend who telephoned me from California. She said: Wo de che gei tou le. I 1) See the famous critique by Lin Yutang, first published in 1930.
understood her to have said 我的车被偷了, and urged her to report it to the police right away. After much confusion, it turns out she was actually speaking bilingually, switching to English for the main verb, i.e., 我的车被偷了. She had parked overnight at a no-parking zone, and the police had towed her car away.

1. Lexical ambiguity

At its basic level ambiguity resides in the word, such as the case of 原因 above, or the oft used example of bank in English, which may refer to the shore of a river or a place for financial transactions. Lexical ambiguity is primarily of two sorts, homophony and polysemy. Homophony arises typically from mergers, a common type of sound change. English words which used to be pronounced with distinct vowels, such as steal and steel, are now homophones because these vowels merged. Their spellings reflect the fact that written language often lags behind spoken language.

The dialects of Chinese vary in their degree of homophony, depending on the mergers which have taken place. A massive merger took place in many northern dialects when the endings of Middle Chinese tone IV, or 入声, were dropped. This is known in Chinese historical phonology as 入声, since the syllables which had these endings merged into the other three tones, i.e., M. C. I, II, and III). However, many southern dialects did not participate in this merger, for instance Cantonese. As an example, 唱, 紅, 和 精 are pronounced with -p, -t, and -k respectively in Cantonese. But in Putonghua, they are all homophones, pronounced sē, with M. C. tone III.

Y. R. Chao once composed an amusing story, tongue in cheek. In Putonghua, the story is read with just one syllable, ‘shi’. There is considerable repetition in the story; for instance, 耳 occurs ten times in this short paragraph. Without the written text, the story is not intelligible. In Cantonese the intelligibility is somewhat improved since a third of the syllables in the story are in M. C. tone IV, and are distinguishable from each other by their -p, -t, and -k endings. The M. C. IV words appear in bold italics below. As examples, 该 ends in -p, 室 ends in -t, and 石 ends in -k.

Here is an amusing dialog centering on the ambiguity of the word 意思, which literally means “meaning”. It takes place after B presents A with a gift, presumably intending to curry favor. The dialog is a miniature study of the pragmatics of Chinese conversations.

A: 你这是什么意思? What is the meaning of this?
B: 没什么意思, 意思意思. It's really nothing; just a small token.
A: 你这就不够意思了. You really shouldn't have.
B: 小意思, 小意思. It's nothing, nothing at all.
A: 你这人真有意思. You are really something.
B: 其实也没有别的意思. Actually it doesn’t mean anything.

3) 赵 1980:149.
4) I thank Dr. LOW Hwee Boon of the Institute of Advanced Studies of the Nanyang Technological University for this interesting dialog (personal communication of February 2011). I have provided the interpretation of the dialog as well as a rough English translation, taking the pragmatics into consideration.

A: 那我就不好意思了。
B: 是我不好意思。

In that case, I (accept and) am embarrassed.
It is I that is embarrassed (with such a humble gift)

1.1. Adaptation

Since languages are complex adaptive systems, one of the ways Putonghua adapted to the massive mergers was to make many words disyllabic, to reduce the numerous instances of homophony. This can be clearly seen by comparing some monosyllabic words in Cantonese with their disyllabic counterparts in Putonghua, which are more redundant.

<table>
<thead>
<tr>
<th>广州话</th>
<th>普通话</th>
<th>广州话</th>
<th>普通话</th>
</tr>
</thead>
<tbody>
<tr>
<td>眼</td>
<td>眼睛</td>
<td>鼻</td>
<td>鼻子</td>
</tr>
<tr>
<td>耳</td>
<td>耳朵</td>
<td>杯</td>
<td>杯子</td>
</tr>
<tr>
<td>女</td>
<td>女儿</td>
<td>被</td>
<td>被子</td>
</tr>
<tr>
<td>观</td>
<td>笔</td>
<td>桌</td>
<td>桌子</td>
</tr>
<tr>
<td>双</td>
<td>翅膀</td>
<td>晃</td>
<td>晃子</td>
</tr>
<tr>
<td>衫</td>
<td>衣服</td>
<td>楣</td>
<td>楣子</td>
</tr>
<tr>
<td>纽</td>
<td>纽扣</td>
<td>袖</td>
<td>袖子</td>
</tr>
<tr>
<td>蟹</td>
<td>螃蟹</td>
<td>褡</td>
<td>褡子</td>
</tr>
<tr>
<td>鳜</td>
<td>鳜鱼</td>
<td>竹</td>
<td>竹子</td>
</tr>
<tr>
<td>有</td>
<td>没有</td>
<td>盛</td>
<td>盛子</td>
</tr>
<tr>
<td>知</td>
<td>知道</td>
<td>蚊</td>
<td>蚊子</td>
</tr>
<tr>
<td>易</td>
<td>容易</td>
<td>仔</td>
<td>儿子</td>
</tr>
</tbody>
</table>

Earlier in the history of the Chinese language, during the 魏晋南北朝

period after the great Han dynasty, a different type of redundancy came into the language in the form of classifiers. This morphological device also helps distinguish words in the modern language. For instance, between ‘lion’ and ‘louse’, which are homophones, i.e., 一头狮子 vs. 一只虱子; or between ‘deer’ and ‘road’, which are homophones, i.e., 一只鹿 vs. 一条路. Even though there is considerable variation in how Chinese dialects use classifiers, nonetheless they are an important linguistic resource for disambiguation.

1.2. Simulation

Nowak et al. (1999) made an interesting start in their quantitative study of homophony by simulating language evolution with virtual agents populating a computer game. They made the intriguing observation that homophones are quite stable across generations. This is because each meaning in the artificial language the agent in the simulation is learning has no competing forms in their simulation.

The simulation approach has been considerably refined in the research of Minett and Gong (2010) in several important ways by taking the context of the word into consideration, rather than treating each word as independent utterances in the work of Nowak et al. They make the crucial distinction between virtual ambiguity, where two homophonous words belong to different morphological classes and thus are unlikely to occur in the same syntactic position, and actual ambiguity, where the two words have similar grammatical distribution and thus compete against each other for what the sentence may mean. Their empirical results confirm the intuition that actual ambiguity is quickly eliminated in the transmission across generations of agents; on the other hand, virtual homophony may persist much longer in evolutionary time because they do not create obstacles in communication.
Another interesting result is their observation that in the simulated sentences, verb homophony is more stable than noun homophony. Their explanation for this result is that transitive sentences contain just one verb but two nouns; hence the nouns tend to compete more, leading to a higher probability for the ambiguity to be resolved. These are all interesting findings toward a more precise understanding of how ambiguity operates in language. We hope that such simulation studies will continue and lead to many more insights on the nature of ambiguity in language.

2. Syntactic ambiguity

When homophonous words occur in an identical context, actual ambiguity occurs. This can be exemplified by the phrase 无法无天, which means literally ‘no law, no heaven’. Homophonous with 发, there is a word 发, which means ‘hair’. The homophonous phrase 无法无天, therefore, can be interpreted to mean literally ‘no hair, no sky’: 天 is polysemous here — it can mean either ‘sky’ in a physical sense, or ‘heaven’ in an abstract sense. In fact, the distinguished writer Edgar Snow may have been tripped up by this ambiguous phrase after an interview with Chairman Mao Zedong, and conjured up an image of a monk under a leaky umbrella to accommodate the second phrase.

A slightly more complicated example also involving ‘hair’ can be found in a Hong Kong advertisement, using a pun to help sell some magic herb. Here the word for ‘hair’, 髪, written in its traditional instead of simplified form, is bracketed in the advertisement. It is homophonous with the word 髮, which combines with 显 to mean ‘discover’. The two interpretations which result are: [1] ‘In a short 30 days, discover a miracle’; and [2] ‘In a short 30 days, hair will appear miraculously’. Punning is a device for deliberate ambiguity that is frequently used to attract attention in diverse linguistic situations.

| 短短 | 30日 | 「髮」現 奇跡の |
| dyun²dyun² | saam¹sap⁶jat⁶ | faa¹jin⁶ kei⁴zik¹ |

2.1 Pronoun Deletion

Even when the constituent words have no ambiguity, they may nonetheless group into sentences which are syntactically ambiguous. This is the case with Y. R. Chao’s famous example of 鸡不吃了, which may mean [1] ‘The chicken is not eating anymore’, or [2] ‘X is not eating any more chicken’, where X represents some deleted subject in the sentence. Interpretation [1] has the default syntactic order, where the subject is 鸡, and the transitive verb 吃 has no object. The verb could take an object, of course, such as 米, ‘rice’.

The default subject in interpretation [2] would be the first person singular pronoun 我, as in 我不吃鸡了. In this case, the object 鸡 is fronted, perhaps for contrastive purposes, as in 我鸡不吃了. (可是再要点饭), i.e., ‘I won’t eat any more chicken, (but I would like a little more rice)’. The ambiguity arises when the subject pronoun 我 is deleted, as is often the case in Chinese discourse, making it possible for 鸡 to be interpreted as the subject.

Similar ambiguity arises in other languages which permit pronoun deletion, such as Italian. For example, the sentence below allows two interpretations. [1] has the interpretation that When Marco calls (him), his...
brother is always happy, where the deleted pronoun refers to his brother. [2] has the interpretation that When Marco calls his brother, (he) is always happy, where the deleted pronoun refers to Marco.

Quando Marco chiama suo fratello è sempre felice.
When Marco calls his brother is always happy.

Often with such cases, the two interpretations can be deliberately distinguished by intonational devices such as pitch and pause. But usually this is not necessary because the context makes it clear which interpretation is intended.

Another aspect of the language which contributes to the ambiguity of 鸡不吃了 is that words like 鸡 and 鱼 can refer either to the animal or to their meat, much as in English ‘chicken’ and ‘fish’. In contrast, words like 牛 and 猪 refer only to the animal; referring to their meat requires the addition of 肉. A similar contrast can be found in English, where terms for animal meat like ‘beef’, ‘pork’ and ‘mutton’ came into the language from Norman French. It is interesting that Chinese and English should share the division among the animals this way, presumably on the basis of some cognitive similarity of the two cultures. In any case, due to this lexical aspect, an utterance like 鸡不吃了 has only the interpretation that corresponds to [1] above, since *我不吃鸡 is not an acceptable sentence in Chinese whereas 我不吃鸡 is fine.

Incidentally, the computer scientist Terry Winograd took advantage of the same ambiguity in the sentence The chickens are ready to eat when he discussed problems of natural language processing. Again, the sentence has the two interpretations discussed above, one with chicken as animal and the other as meat, for essentially the same reasons.

2.2. Indeterminate Attachment

A class of examples of syntactic ambiguity not based on lexical ambiguity has to do with the indeterminate attachment of modifiers. In the Chinese sentence 这两本书我们一起看, two distinct attachments are possible for the modifier 一起. It could attach to the 这两本书, or to the plural pronoun 我们. It could indicate either the books should be read together, side by side, say, or that the readers should be together during the reading. Like the 鸡 example discussed above, the ambiguity arises because the object 这两本书 has been fronted to a position before the modifier. Had the object occurred in its usual post verbal position, 我们一起看两这本书, there would be no such ambiguity.

Similarly in the Japanese sentence, the basic form indicates that Taro chased Hanako. However it is ambiguous because it is indeterminate whether the modifier じてんしゃ-de attaches to Taro or to Hanako. That is, the sentence could mean either that Taro did the chasing by bicycle, or that Hanako escaped by bicycle.

Taro-wa jiteisha-de nigeta Hanako-o oikakemashita.
Taro-subj bicycle-by escaped Hanako-obj chased.

There used to be a favorite joke on the Vaudeville stage, when one comedian would announce that He knows a man with a wooden leg called Peter. The other comedian would then ask him: What is the name of his other leg? The ambiguity here again is where to attach the phrase called Peter. The usual attachment is of course to a man, that the man is called Peter. The joke is to attach it to the immediately preceding noun phrase, a wooden leg, creating the comic effect of wooden legs having names like people.

2.3. Grammatical Classes

Many types of ambiguity are due to cases of homophony across grammatical classes. For the English third person singular pronoun, *his* is in the possessive case, as in *his book*, while *him* is in the objective case, as in *help him*. However, when the pronoun is feminine, the possessive and the objective are homophonous with each other, as in *her book* and *help her*. There is no ambiguity here since the context makes it clear which case is intended.

A similar instance obtains with the word *duck*, typically used as a noun to refer to a farm animal. However, the word has been grammatically extended to a verb, *to duck*, presumably from the head-bopping movement a duck makes when it walks. Similar extensions can be seen in *to wolf down a meal*, to refer to crude manner of eating, and *to ape his gestures*, to refer to imitation. A phrase like *her duck* standing alone is not ambiguous; it will naturally be interpreted as possessive pronoun followed by a noun. However, ambiguity arises in a sentence like *I saw her duck*, when the phrase is embedded and a second interpretation becomes possible where the *her* is the direct object of *saw* and simultaneously the subject of the verb *duck*.

Ambiguities in a sentence frequently arise because it contains two or more elements each of which permits two or more interpretations, due to either homophony or polysemy. Thus in a sentence like *他喜欢炒饭*, the element *喜欢* can either function as an auxiliary verb, such as in *我喜欢*, where *笑* is the main verb; or it can function as the main verb itself, such as in *最喜欢*. The element *炒饭* can either function as a single word, meaning *fried rice*, parallel with *稀饭*, meaning *congee*. Or, it can be segmented into a verb-noun construction, meaning *to fry rice*, parallel with *买书*. Indeed in some dialects, such as in Shanghainese, these two functions of *炒饭* are spoken differently, when tone sandhi operates differently within words and across words. The two interpretations which are possible for *他喜欢炒饭* are therefore: [1] He likes fried rice; and [2] He like to fry rice.

Here is a similar example illustrating the problem of segmentation: *我们没有做不好的事情*. Whereas the *喜欢* in the above sentence can serve either as main verb or as auxiliary verb, the *有* in the present sentence can serve either as main verb or as aspect marker. This polysemy can be seen clearly in a contrasting pair, such as *我没有笔* and *我没有去*. The two interpretations of the present sentence are, respectively: [1] There is nothing that we cannot do well; and [2] We have not done anything bad. In [1], the main verb is *有*, and the element *做不好* is a resultative verb construction, the negative counterpart to *做得好*. In [2], the main verb is *做*, and the *没做* is the negative counterpart to *做了* in Putonghua, which is however *有做* in some dialects, e.g., Minnan.

3. Figurative Language

All languages extend the basic meanings of the words and sentences into various more abstract realms, from the literal to the figurative, idiomatic, and metaphorical. An obvious and very productive area of such extensions is from spatial words, which are more concrete, to temporal words, which are more abstract. Many English words can be used for space as well as for time, e.g., *in the house*, *on the table*, *before the house*, *before Monday*, *long pencil*, *long time*, etc. Similarly in Chinese, *屋内*, *一年内*, *桌上*, *上礼拜*, *年前*, *前天*, *长衫*, *长久*, etc.
A commonly cited example from English is the construction *kicked the bucket* as an idiomatic equivalent to *died*. Whereas the literal construction allows various paraphrasing, the figurative construction is more limited in how they are used. For instance, the passive construction *The bucket was kicked by John* can only be interpreted literally, not figuratively. When such constructions retain their original meaning alongside of the figurative meaning acquired through extension, ambiguity arises. Thus *John kicked the bucket* can simply refer to the physical act, or it can mean that John died.

In normal conversation, we would expect the hearer to interpret the idioms appropriately. However, it appears that people with aphasic disorders, with damage to certain areas of the left hemisphere of the brain, may lose the ability to deal with figurative language. It was found that some Italian aphasic patients when presented with common idioms produce responses which are associated with literal rather than figurative interpretations.

For instance, the idiom *vuotare il sacco* has the literal meaning of *to empty the sack*, but the figurative meaning of *to confess something*. When this idiom was presented to fifteen patients, seven of them produced the correct target response of *Segreto* (secret), while five produced a response triggered by the literal interpretation, i.e., *Zaino* (rucksack), and three patients produced irrelevant responses.

Such failure to retrieve figurative extensions in ambiguous cases is also found in schizophrenic speech. For instance, a patient may understand *court* to refer to the physical grounds where sports like basketball or tennis is played, but fail to understand it as part of a judicial process. In both aphasia and schizophrenia, the semantic network in the brain is compromised, and cases of ambiguity become difficult problems of communication.

It is important that research in areas such as semantic networks in the brain be carried out across languages of different structures embedded in distinct cultures. In recent years, very interesting work has been done in a highly skewed set of cultures, which have been dubbed WEIRD cultures, i.e., Western, Educated, Industrialized, Rich, and Developed. These cultures represent only a small fraction of human populations; studies based on these cultures cannot and should not be extrapolated as information about human language and human mind in general. It is important that research on the languages and cultures of Asia be intensively carried out in these areas of the neurocognition of language so that we can arrive at a more balanced understanding of the human species.

4. Concluding Remarks

To return to the theme of this essay, it is a fascinating question as to why languages should be so ambiguous in the first place. Such ambiguity does not appear to be paralleled in two other areas which require extensive communication via coded messages: one is the area of computer languages, the other is the genetic code. But careful and systematic comparisons from the viewpoint of ambiguity have yet to be made.

As we saw above, ambiguity is built upon homophony and polysemy at the level of words. Clearly, if there were numerous new words to mark finer distinctions, many of the ambiguities would not occur. Suppose we had another word in English, say *chocken*, which means *chicken meat*, in the same way that *beef* means *cow meat*. Then the sentences *The chicken is ready to eat* and *The chocken is ready to eat* are quite distinct, and there would be no ambiguity. Apparently, the English language did not opt for this lexical distinction. Is this because we have a limit to the powers of our lexical

---

14) McKenna and Oh 2005.
memory, as Darwin conjectured?

“We see variability in every tongue, and new words are continually cropping up; but as there is a limit to the powers of memory, single words, like whole languages, gradually become extinct.”

Indeed, the number of distinct sinograms used in the various dynastic histories seems to have remained largely constant, hovering around 8,000. Although there are important differences between the sinogram and the word, Cheng's result is certainly an important step toward elucidating the limits of our lexical memory. This result is consistent with Darwin's conjecture. This is one approach toward explaining the presence of ambiguity in language.

On the other hand, our lexical memory is significantly expanded when we learn additional languages. While there are reports that the vocabulary scores of bilinguals are often lower than those of monolinguals, nonetheless the combined lexicons of multiple languages in a bilingual surely contain many more words than the lexicon of a monolingual. Thus if a language needs to add words like chicken to its lexicon, it would not exceed any limits of lexical memory.

So the above considerations lead us to the following conclusion. Much of ambiguity in language can be eliminated at the expense of adding numerous new words. Adding words would not exceed any limits of lexical memory, though it necessarily increases the complexity of the language. However, by far the majority of ambiguous cases are pre-attentively resolved by linguistic and/or extralinguistic context. Language therefore tolerates its ambiguities, and occasionally exploits them for special purposes.

[References]


Empson, William. 1930. *Seven Types of Ambiguity*.


Minett, J. W. & T. Gong. 2010. Exploring linguistic ambiguity from a simulation perspective. 研究之乐 II：庆祝十五先生七十五寿辰学术论文集, ed. by 潘悟云 & 沈镇伟, 244-64. 上海: 上海教育出版社.


16) Darwin 1871.
17) Sinogram is a translation for 汉字; see Wang and Tsai 2011.


黃國營. 1985. 現代漢語的歧義短語 語言研究 8。

王士元. 2006. 語言是一個複雜適應系統 [Language is a complex adaptive system.] 華大學學報 (哲學社會學版) 21. 5-13。

趙元任. 1980. 語言問題 臺北。