Actus II-- Confusion Philosophy

Having examined the mechanisms of humor, I turn my attention (and yours, too, I hope) to the mechanisms by which linguistic ambiguity may be produced, the way the same sounds may be made to yield more than one meaning at once. The production of linguistic ambiguity has its roots in the perception and processing of speech in general.

2.1 Linguistic Ambiguity

A pun is an **utterance** that relies on an ambiguity for humorous effect; by utterance, I mean any self-contained unit of speech, of any size. Puns can be made on words, phrases, sentences, or even, in extreme cases, groups of several sentences. I specify, though, that a pun must be a **unit of speech**; by linguistic ambiguity, I mean a **verbal** one, one that can be made in speech.

Ambiguity should be distinguished from simple **failure of understanding**. Although ambiguity produces a failure to understand the speaker's meaning, it goes farther: it substitutes an incorrect meaning. In ambiguity, an utterance is misunderstood instead of simply not understood. The difference is well illustrated by a scene from the film **The Right Stuff**, in which Dr. Werner Von Braun presents, to a committee including President Eisenhower and Senator Lyndon Johnson, a plan to put a man in space before the Soviets do. He explains, his German accent contorting his speech, that "**it might be possible to launch a pod**," at which Johnson asks, "**A pot**?!" Von Braun goes on to explain that "**some sort of specimen could be placed inside**.. " and Johnson asks, "**A spaceman?**" Finally, as to what sort of specimen would be placed inside, he "**had in mind a jimp.**" Johnson: "**A jimp? Well, what the hell is that?**" "**A jimpanzee, Senator.**" In the first two cases, there is ambiguity; words are mistaken for other words. In the third, there is only lack of understanding; the words are mistaken for unprocessable nonsense.

The distinction of ambiguity and **vagueness** is fuzzier, for it lies in the amount of information hearers expect to receive. Basically, ambiguous utterances provide all the expected information; they seem completely specified. However, they specify two different meanings. Vague utterances are underspecified; they do not provide enough information, so that there are almost unlimited interpretations possible. The difference is exemplified by the famous oracles to Croesus of Lydia

and Pyrrhus of Epirus. The former was told that if he crossed the Halys River, **a great nation would be destroyed**. This could have been any nation (perhaps some Indian tribe in America; although the practical connection of this would have been unlikely, it would have been permitted linguistically); Croesus took it to be the Persians, while it was meant to be his own. Pyrrhus was supposedly told **Aio te posse Romanos vincere**, "I say that you can conquer the Romans" or "that the Romans can conquer you." (See **3.6** for a discussion of the nature of this ambiguity.) Only two interpretations are possible; the utterance seemed completely specified. But the number of possibilities of the first ("vague") case was reduced by the assumption of relevance to the context; if it were assumed that the oracle's statement had anything to do with the situation, only one of two nations could have been meant. And the second example could have been seen as vague, if hearers always expected, in the neighborhood of the ambiguous subject accusative - infinitive - object accusative construction, further information specifying which noun had which role. It is a matter of what information the hearers are looking for, whether they will be satisfied to see the utterance as vague, just like the Major General in **The Pirates of Penzance** in this exchange with the Pirate King, whose men are about to marry the Major General's daughters rather forcibly.

Maj.: Who, may I ask, are you?

King: We are all single gentlemen.

Maj.: Anything else?

King: No, nothing else.

When utterances are ambiguous, we do not ask, "Anything else?" -- we generally feel we know it all, and find out too late that we do not. The difference of ambiguity and vagueness is thus fairly arbitrary, and, in most cases, whether an utterance is one or the other depends on what the context has suggested and the hearer expects.

There are ambiguities in every form of representation, whether it be mimetic, with an unlimited number of symbols that can be adapted to the subject, (like painting) or symbolic (semiotic) with a limited number of symbols (like language.)[1] I have included several pictures that are visually ambiguous (see

[1] Of course, painting can be said to have a "language," in that it has certain conventions of representation, but it can always be absolutely imitative of life, whereas language itself is

Appendix), and one can imagine ambiguous textures, tastes, smells ("Is that really burnt almond I smell, or did someone just take cyanide?") and non-speech sounds. (An example from training with the Cycling team one morning: "Kirsten, do you have a mouse with you or do you need to oil your chain?") It is important to distinguish jokes that rely on other forms of sensory ambiguities from those that rely on linguistic ambiguity, although sensory ambiguities can be just as coincidental and incongruous. Both require contexts, and knowledge of the multiple possible interpretations of the same indicator, but they are usually more universal, less governed by conventions that apply only to humans or small groups of humans. The line "Is that a rabbit in your pants or are you just glad to see me?"[2] relies on visual ambiguity, while the special use of the same line in the film **Who Framed Roger Rabbit?** in a scene in which a detective enters the presence of a female friend actually concealing the rabbit of the title in his clothes, is a linguistic ambiguity, an application of a familiar phrase out of context, giving it a different meaning. In Act I of **Cyrano de Bergerac** by Edmond Rostand, the hero makes a series of jokes based on visual ambiguities involving his nose, comparing it to a peninsula, a scissors-case, a perch for birds, etc. On the other hand, in Act II, another character (who is supposed to have considerably less wit than Cyrano) makes a series of verbal jokes, interrupting Cyrano's narrative of an epic duel and completing his sentences with appropriate idioms involving the word "nose," e.g., "nose to nose," "stick my nose in," "see no farther than one's nose." Other types of ambiguities can become linguistic, as they can be used in figures of speech which can then be adopted into the lexicon as new meanings for the old words. (E.g., if "bird-perch," were frequently enough used in an imagistic manner as a term for "nose," so that "nose" became another of its standard meanings.) Linguistic ambiguity can require other sensory information to set up the context and the multiple possible interpretations, as in the **Roger Rabbit** example, but it remains linguistic since it is the words themselves that possess multiple meanings.

There are other ways of representing language besides speech, the closest being writing. For most people, in our culture at least, language is a visual

---

absolutely conventional, and cannot imitate life except by arbitrary assignations of sounds to senses.

[2] This is based on Mae West's line, "Is that a gun in your pocket or are you just glad to see me?"

phenomenon as often as an aural one, and the representations of words by letters can be almost as real as those by sound. Our knowledge of spelling can color our pronunciation, our conceptions of words, and our perception of certain types of LAUGHs. On the other hand, we tend to sound out language as we read it, to hear it in our heads. [3] The result is that there are also written puns, which look ambiguous on paper, but when read aloud, can lose all ambiguity or simply not make sense. Or, they can simply work in a way different than that intended or that which an interpretation as visual puns would give. Since puns are possible with only loose phonetic matching, most written puns still work as puns, but will require hearers to tolerate greater changes while still recognizing phonetic similarity, or they may change the type of ambiguity of the pun (as shown in Actus III.) Written puns can always be effective orally if the hearer will stretch his imagination enough, but this stretch often renders them less funny. To look at some examples:

2.1. 1 A bore is a person who tells you a story from a to ZZZZ. (Moger I) "ZZZZ" is the comic-strip symbol for sleep; it is perhaps an imitation of the sound of snoring produced during sleep, but pronounced aloud, it would not sound very much like /zi/, the sound required to fit the pun, since there is no vowel. However, a reader might pronounce the string of "Z"'s exactly that way, as "Zeeeeeee," in which case the pun would still work. It would not work at all in most of the English-speaking world, where the last letter of the alphabet is read as "zed."

2.1. 2 Preferred Taste, a billboard ad for Johnny Walker Red Whisky, must also be seen to be understood, for the crucial letters "red" were set off by being actually printed in that color. One would never hear the word "red" in pronouncing "preferred" unless a strange accentuation were used (as it occasionally is in archaic-sounding poetry.)

2.1. 3 Strum and drang, the description of a banjo concert, relies on the transposition of two letters, and the visual confusion created. ("Is it a typo?" the reader wonders.) Pronounced aloud, the words "strum" and "sturm" resemble each other much less than their visual representations, the difference being in a stressed vowel, which is not so easily overlooked as other types of differences. (see **3.3.c.**)

---

[3] Hammond and Hughes go so far as to assert that "even the sophisticated adult reader forms the words he is reading with his throat muscles." (Chapter 1)

2.1. 4 Dim sum cum laude, subheadline in the Boston **Phoenix** for a review of a Chinese restaurant. This is weak even as a visual pun, since "sum" and "summa," (the word alluded to), have only half their letters in common. Orally, the strong difference in stressed vowel quantity and quality (assuming the non-Latinate pronunciation of *summa*, which seems to be prevalent among those who use the term,) means that there is little connection made.

2.1. 5 Theses are the times that try men's souls. (the author) It works better in writing, as to make the pun only a single letter need be added, rather than an entire syllable which changes the stress. The voicing of the initial sound is also changed. Again, it could be just a misprint.

2.1. 6 How do you tell the difference between a labor organizer and a physical chemist? One is interested in unionization, the other in un-ionization. (David P. Mandell) Without knowledge of spelling, there would be no association of the two words, no ambiguity, and no pun.

Another example is my "**bottomless chiasm**" line on page 12.

Written puns can also be effective to hearers who are highly aware of orthography by requiring a sort of **double reinterpretation**. This is a translation into another language or symbolic system and back, a conception of the written representations that link the sounds that link the different ideas, rather than simply and directly, of the sounds that link the ideas. But in general, though they may obey many of the same rules as verbal puns, fall into the same categories, and work or fail for the analogous reasons, they have much less to do with linguistics than verbal puns, since they require access to level of representation that is not strictly linguistic. In languages whose spelling systems are close to phonetic, such as Latin, written puns are not really a separate category, and the terms "letter" and "sound" can be used almost interchangeably. Identity in appearance means a pretty sure bet of identity of pronunciation. Latin literature contains a large number of anagrams, which could work much better orally than they might in English because of the phoneticity of the writing system. (See Ahl, pp 47-51. One example is the **Aeneid**, lines 8.322-23, explaining the naming of Latium: **LATIUMque vocari/ MALUIT, his quoniam LATUISSET tutus in oris.**) The only non-phonetic aspect of Latin orthography was the length of vowels, which was not indicated. According to Ahl, "However much differences in quantity represent insuperable barriers for the modern scholar, they did not bother the ancients very much." (p55)

Non-phonetic writing systems, such as English, are interesting in that, just as pronunciational similarities, leading to the possibilities of puns, have been created in the spoken language by historical change, so have orthographic similarities been created in the written language. Still, this thesis will be limited to verbal ambiguity, and although many of my examples will have been written originally, I will use them only if they would have a similar effect when pronounced aloud. So, in going through my examples, it is best to read them aloud for the full effect, which will have the added benefit of amusing one's family members or colleagues in faculty lounges.

2.2 Speech Perception and Word Recognition

Although ambiguity is created by speakers in that it comes out of their vocal organs, the hearer plays an equally important role in ambiguity production. To find out what causes utterances to sound like one another, the theories and mechanisms of speech perception must be studied.

Hearers of speech can be compared to modern readers of an ancient document. They are presented with a series of sounds, which, it is hoped will be in their own language. They must then attempt to divide this series up into units, going through all the levels of linguistic representation. This is an incredibly complex task, with various factors making it more complex. Hearers must distinguish non-linguistic noise from the meaningful sounds, just as epigraphers must distinguish letters from smears of ink, blood, or wine. Noise can be mistaken for speech, or it can simply obscure it, as in the film **Blazing Saddles**, when the words of a watchman announcing that the newly arriving "**sheriff is a n-----**" are taken as "**the sheriff is nearer**," due to the ringing of bells at the same time. Such factors as noise, quality of enunciation, speed of speech, and distance, allow ambiguities to be magnified and less similar words to be mistaken for or suggest one another. Hearers, like epigraphers, must next determine what language they are attempting to understand, the natural inclination being to assume that their own language is being spoken. In **Abbot and Costello in the Foreign Legion**, Lou Costello's character's confusion of his officer's response, "**oui**," with the English first person plural pronoun, is played to death. My father told me of American servicemen in Korea singing a popular Japanese song as "**I ain't got no yo-yo,**" (needless to say the actual lyrics had nothing to do with the lack of spool-and-string toys.)

Hearers must then start trying to recognize phones and phonemes, as epigraphers recognizes characters, remembering that just as characters have different forms, whose distribution may be influenced by region, period, or the writer's whim, the hearer must take into account variation by dialect, variation in pronunciation each speaker has, and the variation according to phonetic environment. They must then locate word boundaries, and figure in stress and pitch. They must identify the words, assign them general meanings, attempt to parse them morphologically if they are unfamiliar, and assign literal or figurative meanings. Finally, the words can be taken in relationship to each other and the sentence interpreted, although certain elements, from the reference of anaphoric pronouns to the presence of irony, can only be determined from the surrounding sentences or the hearer/ reader's knowledge of the text and the world. The same process must be repeated for the next sentence, and so on.

The various models of language processing recognize that there are levels of processing (corresponding to the phonetic, phonemic, semantic, and syntactic levels of language, although not all models theorize the same levels) but disagree on the degree to which the several levels can affect each other. This, according to some theories, is where the hearer and the epigrapher part ways. The latter can switch levels arbitrarily; if reading a certain mark on the page would give a certain letter which would give a certain word which would be a certain part of speech requiring a certain interpretation for the sentence, and that interpretation is impossible given his or her assumptions about the text and the world, he or she can go all the way back down to the lowest level, reinterpret the mark, and take it from there all the way back up. He or she has access to every level from every other, and can reinterpret any one to produce change at higher levels until a conclusion, an interpretation consistent with his or her knowledge of the language, the text, and the world is reached. But epigraphers are conscious of what they are doing, and could describe it, state the rules by which they go about it, whereas language processing is to a great extent an unconscious process, and if we knew its rules, there would be no need for linguists (a very sorry state of affairs.) It requires intellectual effort to figure out what we were thinking, to analyze our own speech, to formulate our intuitions into rules.

Thus, puns requiring extensive reanalysis and level-interaction would tend to be more intellectual, less readily digestible to the general hearer, and have less to do with natural properties of language. I call this the **Horseshoe Nail Principle**, that extensive reanalysis up and down levels is difficult and slow, if

not impossible in some cases. It implies that levels of processing can influence adjacent levels, but their influence on other levels is severely constrained. Another important, but related instance in which speakers have access to all levels simultaneously and are consciously aware of the rules is when speaking a foreign language. It can require great intellectual effort, especially at first, and very careful thought given to application of rules. Readers in general (including epigraphers), with a static text before them, also have the opportunity for extensive, multi-level reanalysis, which hearers do not have. They must extract what information they can out of the flow of sounds as it goes by, with no opportunity to reexamine it except based on their memories. It is difficult to perform extensive reanalysis simply because the memory of the exact sounds will fade quickly, leaving only the sense. (This is another reason for separating verbal and written puns.) The humor value of puns involving this sort of extensive analysis will be reduced according to Freud's dictum that overly careful examination defeats pleasure.

The epigrapher and non-native speaker are examples of an **interactive** model, as opposed to an **autonomous** model, in which the levels are compartmentalized, independent of one another. Tyler and Frauenfelder explain how autonomous models of word recognition are strictly "bottom-up" and do not allow "higher sources of knowledge (e.g., lexical, syntactic, semantic, and pragmatic)" to affect lexical recognition. These sources "only contribute to the evaluation and integration of the output of lexical processing, but not to the generation of this output....However, such models permit "lateral" flow of information within a given processing level."(p8) They note further, however, that none of the proposed models is entirely autonomous, nor entirely interactive, but all can be ranked along a continuum in terms of the degree of influence levels can have on one another. Differences in definitions of levels cloud the meaning of the two polar terms even more; is a model like that proposed by J.A. Fodor, in which all linguistic factors form one level with lateral interaction, while all non-linguistic information about the world is excluded from affecting it, interactive or autonomous? The distinction the models make between "sensory input" and "contextual information" is flawed. After all, if sensory input is to be taken to mean the sounds one hears, is it not the context which tells us that a jackhammer is being used in the background, which will affect our perception of sounds? Our recognition of phonetic input, of phones themselves, is conditioned by the surrounding phones. Context cannot be simply previous knowledge; simultaneous

and subsequent information can also have important effects on interpretation. I think it is more correct, and more useful to the analysis of puns, **to see all information received (besides raw acoustic input) as a hierarchy of contexts, which operate analogously no matter what the level.** As for the degree of autonomy of the various linguistic levels, pun evidence tends to support an interactive view, but a constrained one. In other words, all the levels, linguistic and non-, can affect one another, but there are limitations (in accordance with the Horseshoe Nail Principle) on the ease with which this can be done, and they have a part in determining the success of the pun. These constraints will be discussed under the different types of puns.

The other major factor differentiating various word recognition models is the process of **activation** of words, the tapping of them as the appropriate one intended by the speaker. According to the "**logogen**" model proposed by Morton (Pisoni & Luce, p 39) there are certain "passive sensors" representing every word in the lexicon in phonetic, syntactic, and semantic terms, which look for information in the input which matches their stored information. When enough matching input information is received, and a threshold is reached, the logogen is activated. Since "information from any level can combine to push a logogen over its threshold," this is a strongly interactive model. And words can be activated as much by the semantic or syntactic context as by the phonological input, and this is enormously important to puns that involve actual changes of sound. (Those of the First Order, see **3.3**.) The main factor in setting the activation threshold is the knowledge of the frequency of occurrence of the word. Similar is the "**node model**" of Elman and McClellan (Pisoni & Luce, p 43) in which various nodes representing features, phonemes, and words, each with its own activation threshold, are all connected to influence each other, to raise each other to activation, or lower each other from it. "**Cohort theory**," as formulated by W. D. Marslen-Wilson (Pisoni & Luce, p 40) contrasts to logogen theory in proposing that the beginning of the word sets up a "cohort" of possible words. As each phonetic segment is recognized, the choice of possible words is narrowed, but many words are still retained as possibilities until all but one have been eliminated as contrary to the information received. This model is more autonomous, in that recognition is based purely on phonetic input. It originally operated exclusively left-to-right, so that it did not provide for the possibility of rhyming words being mistaken for each other, something on which puns often depend. Later versions, however, abandoned this rigorous all-or-none dropping

out in favor of a more flexible system of various levels of activation, with candidates within the cohort falling or rising in activation with the input. Another model, that of Forster (Pisoni & Luce, p 42) has separate, serial, lexical (where recognition occurs), syntactic (where a structure is constructed), and "message" (where intentions and ideas are decided) processors, with a "general processor" to incorporate real-world knowledge. Activation and the point at which it occurs is important for puns, which require the activation of several words at once.

Most of these models stress **word** recognition. Is the word really the most important unit of meaning in speech? Phonemes may be the basic unit of speech, but they are not the basic unit of meaning, if meaning is taken in its semiotic sense of representing a concept which can represent something in the real world. A particular phoneme does not have the same meaning in every use, although some ancient theories, such as that which Plato has Socrates express in the **Cratylus**, attribute meaning to individual sounds based on the motions of the tongue in forming them: /r/ indicates motion, /l/ slipperiness and smoothness.[1] (He could be right -- is the meaning of a word like **mellifluous** not obvious?) But seriously, folks, if one were asked to describe the meanings of the phoneme /f/, for example, the question would at first seem meaningless. The only way to answer it would be to examine what information the sound is capable of carrying, what distinctions it can make. For instance, in differentiating the minimal pair **fear** and **ear**, /f/ seems to stand for [+ abstract concept, - body part, + feeling, + trepidation, - hearing] and so on. However, in differentiating **fad** and **ad**, for example, it has a completely different meaning, and this does not even include instances when it combines with other phonemes to differentiate meanings. (What is the role of /f/ in differentiating **fruit** and **boot**?) There is a huge number of meanings for the phoneme, but they will never be used more than once.

Morphemes could be seen as basic units of meaning, but again, the same morpheme does not have the same meaning in every case. The morpheme is much more precise than the phoneme, but still very vague. Many morphemes are homophonous, and have a great number of varied meanings, both synchronically and diachronically, and the meaning of a morpheme has a great deal to do with the surrounding morphemes of the word.

But what is a word? I can offer two informal definitions, neither of them without problems: as **(1) a set of phonemes that move around together in**

[1] Snyder (1985) p58-9

syntactic structures and after which a pause might be expected in slow speech, to which a discreet meaning can be assigned out of context, or, as **(2) a semantic unit of meaning, a lexical item, whose actual phonemic form is less important**. The first definition, a rather mechanistic one, would not differentiate homophones, and look at every word by its phonemic representation, as a set of sounds. The "word" [eir] would have the following meanings: to do incorrectly, a gaseous mixture breathed by living things, one who inherits or will inherit, a melody, to bring into the open, etcetera. The homophony, the expression by the same "word," of the second and fifth meanings, which are related diachronically, at least, if not synchronically in speakers' minds, would thus be as coincidental as the homophony of the second and third meanings. The second definition of "word" would allow the various meanings of the set of sounds to be sorted according to speakers' perceptions as to their relatedness, but it runs the risk of including diachronic, etymological knowledge, which speakers are usually not assumed to have. (Also, in a literate culture, the mental representations of different meanings, and the decision as to whether meanings belonged to the same word, would be strongly influenced by spelling.) Etymological knowledge, or beliefs, which among those who do not have the good fortune to have studied historical linguistics are usually incorrect, cannot be ignored, however, especially in the study of wordplay. (It will be treated under Fifth Order Puns, section 3.7.)

But even when it is recognized that the same set of sounds does not mean the same word, it must be recognized that the same word does not always have the same meaning, depending on -- what else?? -- context. In ordinary sentence contexts, words can change in meaning, and when embedded in set expressions, in idioms, they can take on entirely new, specific meanings which seem to be little related to the original meaning, and can be inseparable from the rest of the expression. Therefore, should fixed expressions be counted as words? They behave in many ways like lexical items. They are made up of words just as words are made up of morphemes, and in both cases the smaller units lose their individual meaning in the larger whole, which is not analyzed. We hear "return," not "re-" + "turn," and when we hear "come a cropper," we do not ask what a "cropper" is. We do not ask, that is, if we are familiar with the expression; only when the meaning of the whole unit is not known do we analyze on lower levels, a pursuit which is often not very useful. Could we really recover the meaning of either of the examples by analysis? It is thus necessary to expand the definition of the **lexicon** to include idioms which may not be semantically or even syntactically

analyzable. The various types of puns will show that idioms behave exactly like words in puns, and that the all-important factor is the hearer's previous knowledge of the lexical items. Morphemes can often behave like lexical items as well, again depending on the hearer's knowledge.

The smallest unit of meaning should be the largest memorized unit which is activated as one, the largest unit which does not have to be analyzed to obtain its meaning. Words and idioms are the largest units for which we have a prestored mental representation to which new information is matched; for larger units, such as sentences, we have only rules for breaking them down and analyzing them. Some units with sentential properties (i.e., subjects, predicates, etc.) do not need to be analyzed; set phrases can often have full sentence structure (or even multiple sentence structure: one could imagine a minister for whom the marriage formula has become meaningless, unanalyzed; I have examples below of formulae which children memorize without parsing, and in high school, those of us who studied Russian would teach lines from dialogues to other students, who would memorize them as undifferentiated stream of sounds.) Not all words or idioms can get by without analysis, only familiar ones, and the analysis of an unfamiliar word or idiom into its component morphemes or words will be either a largely conscious, intellectual process (with words) or a rather useless one (in both cases.) **Different levels involve different types of processing.** Phonemes need only be **recognized**, not **interpreted**; lexical items, as meaningful units, need to be recognized and interpreted; most sentences need to be **analyzed**. Still, as I hope to show by the examples in the various orders of puns in Actus III, **every level functions essentially like every other in using units which of themselves have no meaning, or vague meaning, and whose full meanings are thus produced by context, on the same level or from other levels.**

2.3 Role and Types of Contexts

As Hardy Hansen and Gerald M. Quinn put it, (all too often) in their **Greek: An Intensive Course,** in distinguishing identical forms (of verbs, in this case), "Context will determine meaning." To help in the task of differentiating apparently similar units of speech, hearers have two aids: knowledge of the language and knowledge of the world. Let us assume that as native speakers of their language, they know it perfectly; thus they can recognize every sound, know the meaning of every word, parse every grammatical construction. Of course, there will always be limits to this knowledge, such as unknown lexical items,

additional meanings, or turns of phrase; extreme dialect variation within a language would also come under this heading. Also, we do not just know lexical items, but their syntactic and semantic subcategorizations, and how frequently they are used. Frequency plays a role in most models of word recognition; in the logogen theory, for instance, it determines each logogen's activation threshold, so that more common words are activated more easily. Pritchett (1987, p141ff) cites Ford, Bresnan and Kaplan's ideas of **lexical preference** and **strongest lexical form** (**SLF**), a subcategorization frame for each word that dictates the syntactic and semantic function it is **most likely** to have. However, these ideas do not seem to be context-sensitive; although a word may have a strongest lexical form that is statistically or conceptually its most usual one, different contexts (semantic, syntactic, or situational) can make different forms the strongest, without the usual SLF even being thought of. Puns often set up a situation in which a less common meaning is the obvious one, then use the word in its most usual sense as well, which in the new context sounds bizarre and humorous. (For example, in **L'Odysée d'Astérix,** (p33) Astérix and Obélix are waiting at the foot of the walls of a city, trying to sneak in, when they are attacked by a Roman patrol. Obélix cries out, "**Ils veulent me mettre au pied du mur!**""They want to put me at the foot of the wall!" = "They want us to have our backs to the wall, have no escape." Since cities are no longer walled, the figurative sense is the more common one, except in this instance.) Language knowledge is basically unconscious; we cannot necessarily say what we know about language, we just use it. We reject sentences as bad without quite knowing why sometimes, even if we are linguists.

To fill the gaps in knowledge, and also those which knowledge cannot cover, requires knowledge of the world, knowing how things and ideas work rather than just words. Knowledge of the language would include everything in the utterance that is not the specific unit up for interpretation; if the unit is a sound, it is the surrounding sounds; if a morpheme, the surrounding morphemes and the word itself, if a word, the surrounding words (but only within a certain limit. Material fifteen minutes later would be unlikely to affect it.) Just as Chomsky's grammar originally could account for most grammatical sentences, but would still produce some which made no sense (such as "Colorless green ideas sleep furiously,"(Raskin, p50)) knowledge of the language can only go so far in deciding whether a sentence would be produced by a native speaker. It is the hearer's knowledge of the world which would make the example incomprehensible, even though meanings and grammatical roles could be

assigned to the words. World-knowledge includes everything that describes the situation in which the sounds are heard, from the relationship of speaker and hearer, the literary setting, the events of the moment, the characteristics of speech of the parties, the language spoken, the narrations used to set up a joke, or the outside noises.

The line between knowledge of language and knowledge of the world is a fuzzy one, for linguists are constantly trying to push the outside of the envelope of linguistics, adding the "psycho-" and "socio-" prefixes to include more and more of the world. For certain sentences, it is clear, it takes no knowledge other than that of language to reject them. The noise of a machine gun or the song of a bird are obviously not processable to us as language, nor is the string "**geemeeahl yehmani**," nor "**a synemast florted barioculously**." But since language describes the world, the unfamiliarity of a lexical item may be attributable to the lack of knowledge of the thing itself or to that of the word for it. When certain words cannot be used together, it can be for syntactic reasons, or that the concepts are alien to each other. "I went the store," is not grammatical, whereas "I visited the store" is, simply because the verb "go" requires a preposition. On the other hand, "I went to the Existentialism," while grammatical, does not make any sense if one knows that the verb "go" requires a concrete object. (This is the fuzzy zone.) But what about "I went to President Reagan's funeral."? This sentence is impossible, as it describes a non-existent event, (fails to refer, in speech act terms) and might be rejected because of the hearer's knowledge of the world. Both knowledge of the world and knowledge of the language can be used to disambiguate ambiguous situations, to reject certain alternatives and come down to the intended one. However, both of them are subject to being stretched.

When the hearers have some reason to attempt to find a meaning in what appears to be nonsense, they will do so. If they expect to find a meaning, or feel they are expected to, or if their minds are bent in that direction, they might try to, and succeed. Each of the examples in the previous paragraph might have an interpretation assigned to it if one were looked for, if the sentences were not dismissed. In the first case, one might find some group of sounds of one's language that were suggested by the noise, as I once saw a soap opera character hear the noise of a cuckoo clock as taunting him as a "cuckold." The second example, heard from behind on a dark street, might be taken as "Give me all your money," or, at the U.N., as the name of a diplomat. In the third example, morphemes can be identified, and grammatical roles assigned, and even some

guess made as to the words' meanings. With enough imagination, the other "ungrammatical" sentences above could also be made interpretable. One could find another interpretation for "existentialism," as the name of a course, or a bar, for example, making that sentence make sense. In a science-fiction novel, it might be perfectly logical for a character to attend Reagan's funeral. If one tried hard enough, one could even invent a situation, a context, in which "colorless green ideas sleep furiously" made sense. In context, anything can be meaningful.

When the context suggests that what seems like nonsense really has some structure, the hearer will do his best to interpret it, which is the principal upon which much of James Joyce's work works. And since context can cause one to find meanings where none are intended, by extension, it can cause one to find meanings other than the intended ones in meaningful language. Hearers tend to interpret all new information in the context of theirselves and the situation and outlook of the moment. When I worked as a bicycle courier, I tended to interpret everything in terms relevant to my occupation, so that a song of which the lyrics ran "**I am the passenger/ I ride and ride**," (Iggy Pop, "The Passenger,") I heard as "**I am the messenger**," and one which went "**I am the night warrior**" I heard as "**I am the bike courier.**" (Jerry Harrison, "Cherokee Chief.") Listening recently to a song which I had heard several times before, I was struck by what I was sure was my own last name, when I had always heard the word "story" before. (Popular music is especially good for this sort of ambiguity finding, which can resemble a Rorschach inkblot test in its dependence on the psychological state and associations of the hearer, as there is always noise --er, music, I mean, obscuring the language. Famous Minnesotan "Weird Al" Yankovic has made a career of parodying other people's songs, usually by distorting the key, repeated phrase of the original song and then building his song around it. Thus, "**I Love Rock and Roll**," by Joan Jett and the Blackhearts, becomes "**I Love Rocky Road**," a song entirely about ice cream. Michael Jackson's "**Beat It**" becomes "**Eat It**," entirely about food. Hearing "Weird Al"'s versions for the first time, it can be a while before one recognizes that the lyrics are not the familiar ones, since one probably never really understood them in the first place.)

Outside of its context, on the other hand, meaning can deteriorate into nonsense. Ask a child to carefully recite the Pledge of Allegiance, and he or she might produce something like "**I led the pigeons to the flag...and to the republic**

for Richard Stans, one naked individual, under guard..."[3] In church, he or she might sing hymns about the famous spectacle-needing grizzly, "**Gladly, the cross-eyed bear**", or promise to "**follow Henry Joyce**." The child does not understand the patriotic or religious context, and the words are empty, devoid of meaning, just something to recite. There is no reason not to substitute the names of familiar objects or activities for unfamiliar ones. The situation is even worse with foreign languages; children will produce creatively mangled versions of "Frère Jacques," and I remember interpreting the lyrics of the Beatles' "Michelle, ma Belle," "**Ce sont des mots qui vont très bien ensemble**," as "**Sunday, Monday, you play piano some**."

2.4 Tolerance of Variation

Ambiguities are often the result of imperfect hearing or the willingness to hear imperfectly, to overlook certain information. One good reason in general for not dismissing nonsense, i.e., non-grammatical (in the widest sense) speech, is that it might simply contain some unintended error, or the hearer may have made some error, and knows it. Once the error has been located and corrected in the hearer's mind, the sentence can be correctly interpreted. But we need a certain looseness, a flexibility to the rules, since people are not always going to speak exactly in accordance with them (when they make errors) or they may not possess the rules, if they are children or non-native speakers. In phonetics, "the mapping between phonemes and their allophonic variations is highly variable and extremely ambiguous." (Church,p 54) There is great variation in individual speakers within phonemes; phones that are assigned to one phoneme one moment must be assigned to another the next. (Sometimes, though, this phonetic variation can serve to indicate morpheme, word, and phrase boundaries.[1]) Speakers of different dialects will also draw phoneme boundaries differently, and use words differently. Speakers may even have impediments. Although "any variation will obscure or complicate the relation between the input and prestored lexical representations,"[2] most can be readily overlooked. Just as the epigrapher must recognize the possibility of use of different dialectical forms as well as scribal errors, the hearer must allow for variation and errors. The question is, is this

---

[3] Examples are from William Safire's "On Language" column.

[1] Bradley and Forster, p116

[2] Frazier II, p158

process conscious or unconscious? According to Bradley and Forster, participants in experiments often do not report errors, reporting the correct forms instead, having automatically adjusted for them. (p110) It is a matter of degree, of the size of the error or variation, and the familiarity with that type of it. If one started a job at a day-care center, at first each ungrammatical sentence uttered by a child would need to be consciously analyzed, but as one became accustomed to errors, one might not even notice them. The same would happen if an Englishman moved to Jamaica; he would first have to translate the accent (of Jamaican English; the differences in lexicon and syntax of Jamaican Creole would be another story) consciously into his own sounds, but soon would do it automatically. The important thing is that we learn not to be too demanding in hearing. In **3.3** we will see how puns can work exactly like dialect variation and speech errors. Because we allow this sort of variation, we do not immediately throw out as ungrammatical the variation required by puns, and are thus tricked by them.

Another reason why words and other units have a certain "natural ambiguity" to hearers is that they do not hear the entire word before making an identification. "Under normal listening conditions, a human listener may not have to identify all the phonetic input to recognize the word in an utterance." (Pisoni & Luce, p 38) We recognize sounds based on only a few features, words based on only a few phonemes, phrases and sentences based only on a few words; our ears and processing equipment are basically lazy, and do no more work than seems necessary. This can be seen more positively as a matter of efficiency; why go through the trouble of recognizing all the input when only part of it is necessary to understand the sense of the utterance? Certain information will play the key role, have the most attention paid to it, for various reasons. Stressed syllables and words naturally are given more weight than non-stressed ones; it is they alone which are used to initiate a lexical search, according to Grosjean and Gee. (p144) These two further give an entire hierarchy of the units to which will tend to be recognized individually, with stressed words carrying new information at the top, and unstressed syllables in polysyllabic words at the bottom. (A sort of best-stressed list.) The less information a unit carries, the less stress on it, and the less likely it is to be noticed, so the more easily it can be changed for punning purposes. W.D. Marslen-Wilson performed various experiments showing that words are recognized in about the same time no matter where in them the distinctive element appears, and that, especially in context, words are "selected early." (p76) The same applies to expressions, phrases, and, as shown by so-called

"garden path sentences," to syntactic constructions in sentences; based on earlier information, we fill in the later information and identify the unit. "You do not necessarily retrieve all the semantic information that you possess about a word. If you lack some information, the gap may go unnoticed where it is not crucial to understanding the sentence," (Johnson-Laird, p200); I would emend this to, "when it does not *seem* to be crucial." Puns use information that might have been ignored as irrelevant to create ambiguities. Early and incomplete selection allows hearers in ordinary conversation still to deal with and interpret input that is fragmented and incomplete, whether because we have not heard all the sounds or because we do not know all the words. Limited information leads us to expect the rest; jokes often use this play on expectation with situations; puns do it with words. My favorite example is the song that goes "**My baby fell out of the window/ You probably think her head split/ But luck was with her that morning/ And she fell in a bucket of shhhhh------aving cream!"**

In hearing language, then, hearers are used to receiving input that does not match their representations and rules exactly. They construct their interpretations based on assumptions and probabilities. Using incomplete information, and allowing for the possibility of violations of their rules and knowledge, they then try to come up with a most likely interpretation, one which fits the data and the rules best, by whatever criteria they are judging, like the scientist who tries to fit a line or a curve to the data on his graph, discarding some of it, assigning some of it less weight than the rest. This is the method of epigraphers; if they must ignore a letter or a sound, or even a word or a deviant form, they do so, and seek an explanation or a justification in scribal error or basic stupidity, an explanation which hearers also use, along with slips of the tongue or ear or any other factor that could plausibly explain deviation from the linguistic norm. Epigraphers, in trying to reconstruct a text, assume that it follows their rules of grammar, but also assume that what makes sense to them is the intended meaning. They must weigh the two considerations against one another, for sometimes, to interpret by the rules would produce a meaningless sentence. They must go back and forth, up and down the levels; if the text makes no sense, the lexical meaning of one word might be changed, or its role in the sentence; this might require accepting a change in its spelling. (Unlike hearers, of course, they are not limited by the Horseshoe Nail Principle.) The other possibility is to change their assumptions about what makes sense, that is, to accept that the text is saying something different from what they originally expected it to say. At every level, there is a

context of what is reasonable; the same applies to speech. Hearers can overlook mispronunciations and grammatical errors and still understand the speaker; they can hear only a few sounds or a few syllables and still reconstruct the gist of the utterance. They may even realize that they heard the speaker correctly and that their assumptions about the world and language are wrong. As Sherlock Holmes said, when one has eliminated everything that is possible, whatever is left, no matter how improbable, must be the answer.

This probabilistic model, although it may apply to the epigrapher or the conscious hearer, and well as to lexical activation, is problematical when applied to syntactic processing. It is not so much a question of consciousness as of recognition versus processing. According to Lyn Frazier[1], "perceivers seem systematically to use syntactic well-formedness conditions in processing sentences. They apparently do not rely on a collection of probabilistic heuristics derived from the grammar." (p31) Yet according to Bond and Garnes, (p129) "Speech perception employs heuristic strategies as opposed to formal grammatical operations." Pritchett (l987) discusses various strategies, built into the processor, which dictate why we process garden path sentences the way we do, but they are not probabilistic factors to be weighed against one another, but short-cuts and assumptions. They seem like automatic, almost mechanistic reflexes which are easily tripped up in unconventional contexts. Pritchett's final conclusion is that (290) unprocessability results from "local violations which necessitate reanalysis beyond the bounds of the parser." Syntactic reanalysis thus seems to be rather more difficult than reassignment of meanings to words. Since puns depend on reassignment and to a lesser extent, reanalysis induced by context, this fact contributes to some extent to the appreciation of puns or lack thereof.

As stated at the beginning of this chapter, linguistic ambiguity, deliberately created for punning, has its root in speech processing. The way we hear and process language leads us by necessity to be sloppy and loose, to accept variation from rules, to be flexible. The same flexibility that allows us to understand language allows us to misunderstand it, as we may jump to conclusions about an utterance based on little information about it. Ambiguity, especially humorous ambiguity, depends not only on a **lack of attention** but on a **suspension of disbelief,** and the two are connected. Hearers of normal speech, intended to convey meaning, are willing to overlook divergence from rules in order to

[1] From the Lewis Carroll beast who "always looks grave at a pun."

understand the speech and get at the meaning; these same people can be suckered by the unscrupulous punster into initially accepting his or her less-than-totally-meaningful (but often not totally unmeaningful) speech as normal, and the variations in it, which lead to ambiguity, as normal and unintended. Only too late do the hearers/victims realize the truth. Or, when there is a greater awareness that the pun is being made, it still depends on the suspension of disbelief by the hearers, their willingness to overlook changes in sound, the admission that boundaries and distinctions can be fuzzy, an allowance which they have developed in order to cope with variation and speech errors. Puns thus work best when they are parallel to actual variation and speech and hearing errors, with which hearers are accustomed to dealing. The descriptions of speech errors in several articles (Fromkin, Bond & Garnes) match very closely to puns, and the specific relationships will be examined in the next chapter.

The requirement of suspension of disbelief has two consequences, however, corresponding to the two extreme degrees to which it can be taken. Except in rare cases, if the hearers wanted to be particular, or were desperate to hear only one meaning, they could insist that only one was possible, that the others were ruled out by context or by the rules of normal usage. (This same hearer, the "snark,"[2] or "language fascist," would be unable to understand dialects other than his own, not to mention metaphor. Computerized speech processing systems often act this way.)

Or, if there is too much variation, the suspension of disbelief can be strained. It is similar to another common situation requiring suspension of disbelief, the watching of an adventure film. In **Raiders of the Lost Ark,** when Indiana Jones climbs onto a German U-boat in the middle of the Mediterranean and either rides the outside hundreds of miles underwater or sneaks inside, where there would be no room to hide, our willingness to suspend disbelief is strained, overused by the screenplay. By the film's own rules, Jones is an ordinary guy, if very lucky, not a superman. If he can swim underwater for hundreds of miles or whatever, then we cannot sympathize with him anymore; there is no struggle, anything is possible. We may even laugh at this point. Puns pass themselves off as reasonable within the language; when the cease to be so, they lose much of their appeal. There is a limited amount of variation which can be tolerated and enjoyed, but when a certain limit is exceeded, the hearer's enjoyment is reduced. This

[2]After the monster In the Carroll poem who "Always looks grave at a pun."

constraint is best expressed by **Hale's Law**. Entering the Linguistics Department one day, I met two distinguished members of it and inquired of them if they knew the whereabouts of my advisor, and as I walked away, one of the professors responded, at least as I thought I heard it, that my advisor had "**gone out for some wine and cheese**." As this was rather out of context, I asked the professor to repeat himself and was told that the actual utterance was "**gone out for some lunch**." The other professor remarked of my mistake, "With homophones like those, anything is a pun." Exactly true; and as W.S. Gilbert wrote in **The Gondoliers**, (Act II) "When everyone is somebody/ Then no one's anybody." (This is **Gilbert's Corollary**.) When everything is a pun, nothing is; they cease to become funny; there are too many of them, one does not no where to begin. There is no sense of unexpectedness, of recognizing the familiar in an unlikely location. There must be certain constraints, limits on substitution and variation, in order to keep puns funny.

The degree of constraint from Hale's Law will vary for each person. And the law can be expressed in terms of word recognition. When expressions such as "He went out for some wine and cheese" and "He went out for some lunch" cause activation of each other, it means that the activation thresholds of lexical items are very low. Almost anything will activate them. On the other hand, assuming normal levels of activation thresholds, then words which differ significantly in sound will not activate each other. Puns work by activating more than one word at once, and they can do this since activation thresholds do not require total specification or all the information. If they do not even match this template in the limited way required to activate two words, then there will be no reason to confuse them. The pun will not succeed. Specific applications of Hale's Law and the constraints it produces will be discussed in the next chapter.