Ling 115 **"A corpus-linguistic analysis of English *–ic* and *-ical* adjectives"**
Kelli Wiseth October 27, 2012

What research questions does the article address?
Gries takes an in-depth look at English adjectives formed using the suffixes *–ic* and *–ical*, with respect to questions of synonymy. His main research questions concern adjective pairs (-ic/-ical) in the realm of semantics (meaning):

1) Are such adjectives synonyms, and if not, how can they be differentiated?

2) Do the suffixes –ic and –ical in and of themselves contain a consistent (and different) "meaning component" that can be used to distinguish between the pairs that use both?

Gries provides a fair amount of background on previous studies that have attempted to isolate the semantic differences between –ic/-ical adjectives. As a starting point, he extends Marsden's "forced choice"-study results to identify the top 15 most common such –ic/-ical adjectives, sorted from "most easily distinguishable" to "least easily distinguishes" (per the various dictionary definitions and other reference materials).

He looks at several specific examples from the adjective pairs in terms of previous researchers' theories, which represent thinking from a few different general schools of thought on the subject. At one end of the spectrum, we have those who are "pessimistic about whether there's a discernible [rule-based] pattern, including Fowler, Snell, Ross.

At the other end of the spectrum, we have those who seem to have identified "tendencies" (rather than rules). Proposals from this end of the spectrum include suggesting that distinctions can be drawn based upon whether the adjective derives from "being more directly connected to the root idea of the [adjective's] meaning," and that –ic adjectives are used more in scientific fields, and '-ic being more specific, while –ical is more general,' (Marchand, Fournier).

Gries looks at the data in the context of the various theories and, for the most part, finds the flaws in logic or study design. The bottom line is that no single method for distinguishing meaning between –ic/-ical adjectives succeeds in every instance (every root adjective).

Why do you find the research question interesting?
I found this research question interesting because the contrast between –ic/-ical is an excellent example of the perplexing nature of the English language that is often not helped by 'prescriptive' grammarians nor by use of dictionaries, which, as Gries points out, often do not use the same criteria even within their own definitions.

As discussed in class, even native English speakers have no sense of "rules" that can be easily applied as to when to choose one or the other, so for non-native speakers trying to learn the language, or trying to simply improve their speaking or writing skills, it must be a nightmare. (As a native speaker of English, I know that I often rely simply on "how it sounds," whether –ic sounds better with one word or another, or if –ical sounds better with the same word. But that approach can't be used to guide a learner.

Thus, Gries' research questions are tied to a very practical need—to be able to understand at a non-subjective level how these words are actually represented; when their meanings actually do overlap (and in what contexts); and when they do not.

What conclusion about the question does the article draw?
Gries draws several conclusions and presents many ancillary findings throughout the paper, but his main points in conclusion seem to be:

Recapitulating his use of Tversky's 'contrast model' and Biber's R1 collolations work as the two techniques that he leveraged to come up with a means to assess both the degrees similarity and the degrees of difference (as two countervailing phenomenon), the results of which prove that:

1) many so-called 'synonyms' (logistic(al)/symetric(al))actually pattern do not in fact pattern the same, and suggests that dictionaries be revised to reflect the facts of usage

2) corpus collection and inspection techniques should identify discriminating collocates

Gries restates that discriminating collocates techniques provide "new distinctions that differentiate between adjectives that sometimes defy easy characterization in terms of significant collocates," (examples magic(al), numeric(al)), although our class discussion seemed to disagree with this conclusion.

He also touts the advantages of the corpus-based methodology, in general, that is, more "natural (than ... dictionaries), ...more representative," objectively and statistically gathered, and avoids "distortion caused by prescriptive attitudes."

Gries also restates his use of the corpus-based approach to validate earlier findings by Marchand and Kaunisto, both of which were discussed throughout the paper and to which Gries applies his corpus and statistical techniques. As you know, Kaunisto's 'economy principal,' suggests that in the 'interests of economy, shorter is better,' and so the –ic form of the adjective should occur with greater frequency in those cases where it includes hyphenated linguistic material. If two "adjectives do not make much of a difference in terms of semantics, then there is no reason not to have economical derivational processes, but... if the two adjectives are semantically very different," then the economy principal can't override the choice. Gries' example of "symetric multi-processing" does a good job of proving this latter point.

Do you find the authors' approach satisfactory? If not, how else would you do it?
What I especially liked about Gries approach (more specifically, the paper itself) in addition to the preliminary background of previous studies) was his detailed discussion of Tversky's 'model of similarity'[1] and his discussion of Biber's technique "to identify different senses of polysemous words using R1 collocates," and how both of these provided a basis for his corpus-based 'estimation of significant collocate overlap" (ESCO) which takes into account not only the significant R1 collocates shared by an adjective pair, but also the significant collocates not shared.

[1] In class you mentioned Tversky's similarity model is often used in the realm of phonetics, and Gries discussion in fact reminded me of Pierrehumbert's paper, so I took another look at that paper and see that Pierrehumbert does reference Tversky's 'weighting method' as a technique (paraphrase: [that could have been applied in her study, to capture additional observed regularities that were otherwise insufficiently distinguishable]. Tversky's theory seems like something worth taking a closer look at one of these days.

I thought his use of graphics to present the complex inter-relationships among some of the result sets was helpful, particularly, the $ESCO_2$ chart that conveys something of a snapshot of his findings, as well as Tversky-specific renderings (Figure 2 and 3) that highlight collocates with respect to the symmetry of two specific adjective pairs (logistic(al); symetric(al). These helped in understanding of course the discussion, which was fairly comprehensive and in-depth.

The details of Gries' adaption of his study to claims of Marchand and Kaunisto was also informative and educational (in the sense that we see how current research can shed light on previous research, as well as how various statistical methods can be applied).

Finally, I thought your follow-up discussion (re: politic/political) was very helpful, enabling us to extract actual data using techniques we've been learning to have some basis for comparison with this study's results. Ideally, we should be able to apply the various statistical techniques discussed in Gries (and all the other papers we're reading) to actual data, and this gives us that opportunity (although it may take me some time to get that done! ;-)