The English comparative - Phonology and Usage

In English, the comparative can be formed either synthetically (*prouder*) or analytically (*more proud*). Quirk *et al*. (1985: 461) identify word length as the major factor in the distribution, suggesting that the synthetic comparative is formed by monosyllabic adjectives, while trisyllabic or longer adjectives form the analytic comparative. Disyllabic adjectives are said to exhibit variation. While word length does seem to have a strong effect, exceptions such as *more apt* or *trustworthier* indicate that more factors have to be considered.

Mondorf (2003) summarizes a number of phonological factors that influence the alternation. First, the analytic comparative is favored if the *-er* suffix serves as a buffer between two stressed syllables, as in *a frésher sálad*. Second, adjectives ending in /r/ such as *austere* tend to form the analytic comparative, suggesting that the similarity between the final segment and the *-er* suffix disfavors the synthetic variant. Third, adjectives ending in a consonant cluster overwhelmingly form the analytic comparative, such that *more apt* is preferred to *apter*.

The present paper aims to quantify the relative influence of factors in the formation of the English comparative. To this end, all English adjectives that alternate between the synthetic and analytic comparative are identified. Relying on data from the Google n-gram corpus (Brants and Franz 2006), 730 adjectives are retrieved that enter the alternation. For each adjective, it is determined how often it occurs in its positive form as well as in the synthetic and analytic comparative. Four examples are shown in Table 1 (next page).

From the frequencies collected, it can be determined what ratio of the comparative forms of a given adjective are formed synthetically. As is shown, *apt* only sporadically forms the synthetic comparative, while *early* and *simple* occur almost exclusively with this variant. Elements such as *true* do not exhibit a bias of such strength. Figure 1 (next page) shows that the alternating adjectives are very evenly distributed across the synthetic and analytic comparative.

Each adjective in the database is coded for phonological properties that have been discussed in the literature as factors influencing the comparative alternation. Among these are the number of syllables, the stress pattern, and characteristics of the final segment. All of them are entered into a multiple regression in order to determine which factors are predictive of the ratio of synthetic and analytic comparatives. A low R-squared (.132) indicates that the phonological form of an adjective alone does not strongly determine how it forms the comparative.

In a second multiple regression, frequency values are also entered as factors. The resulting R-squared is much higher (.403). Besides phonological factors such as the number of syllables, final stress, final /i/, and whether the *-er* suffix adds a syllable to the word or not (cf. *simple-simpler*), the absolute frequency of the analytic comparative strongly predicts the ratio of analytic and synthetic comparatives (cf. Table 2, next page). This suggests a usage effect: If speakers have heard an adjective being used in the analytic comparative (i.e. *more glad*), they are likely to repeat this usage, even if the phonological characteristics of the adjectives would make it a perfect candidate for the synthetic comparative.

Table 1: Frequency data for four alternating adjectives

ADJECTIVE	POSITIVE	ANALYT. COMP	SYNTH. COMP	ANALYT/SYNTH RATIO
apt	6,285,854	6391	128	49.929
early	100,104,379	11,558	1,333,629	0.008
simple	85,839,920	22,974	265,075	0.087
true	96,531,418	60,051	35,844	1.675

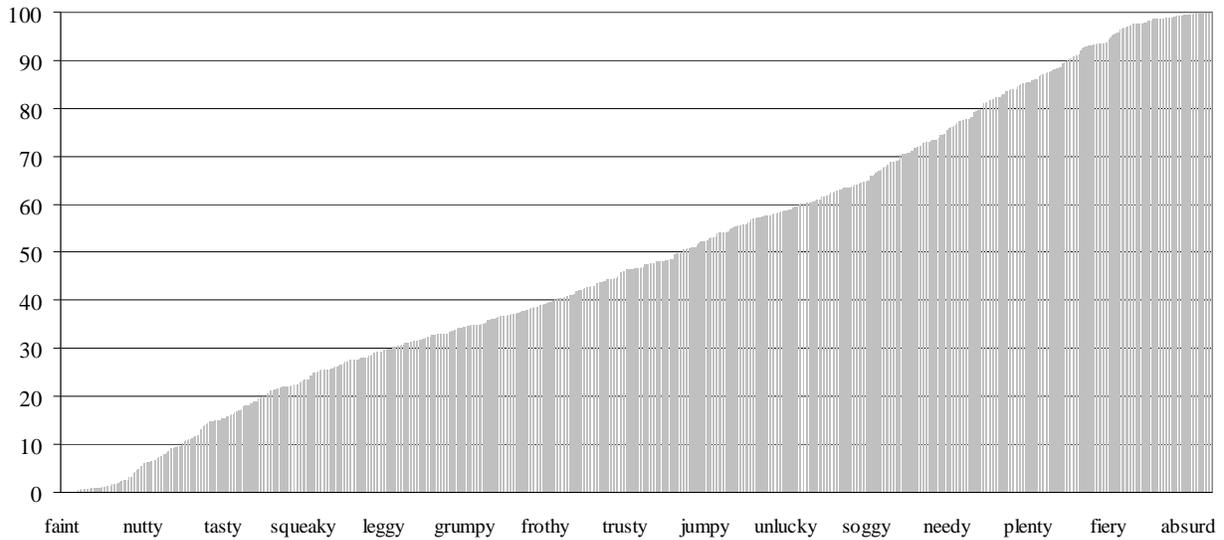Figure 1: Observed percentages of analytic comparatives

Table 2: Factors predicting a high analytic/synthetic ratio

FACTOR	BETA	T	SIG
Absolute frequency of analytic comparative	.545	18.282	.000
High number of syllables	.490	7.472	.000
Synthetic comparative has one more syllable than positive	.157	4.834	.000
Final segment is /i/	.199	3.394	.001
Adjective has word-final stress	.252	3.016	.003

References

Brants, Thorsten and Alex Franz. 2006. *Web 1T 5-gram Version 1*. Philadelphia: Linguistic Data Consortium.

Mondorf, Britta. 2003. Support for More-Support. In G. Rohdenburg and B. Mondorf (eds), *Determinants of grammatical variation in English*. Berlin: Mouton de Gruyter, 251-304.

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey and Jan Svartvik. 1985. *A comprehensive grammar of the English language.* New York: Longman.