

Kansas Working Papers
in
Linguistics

edited by

Kumiko Ichihashi
Mary Sarah Linn

Partial funding for this journal is provided by the
Graduate Student Council for the Student Activity Fee.

© Linguistics Graduate Student Association
University of Kansas, 1991

Volume 16
1991

Kansas Working Papers in Linguistics
Volume 16

Part I: General Linguistics

Collaboration on Topic Change in Conversation Mary Howe	1
Stories in Conversation Roberta Senner Hofer	15
It's Like, 'What's Happening in the Evolution of Like?': A Theory of Grammaticalization Teresa Meehan	37
Folk Etymology (in English and Elsewhere) Adrian Poruciuc	53
Liquid Assimilation in Havana Spanish Fenfang Hwa	59
The Syntax-Phonology Interface as the Key to Metricality: Evidence from Taiwanese Folk Songs Yuchau E. Hsiao	77
Causative Constructions in Nepali Kaoru Horie	99

Part II: Studies in Native American Languages

Transitivity Indicators -Yu and -Wi in Hualapai Kumiko Ichihashi	111
Proto-Algic III:..Pronouns Paul Proulx	129
<i>Cumulative Contents to Volumes 1-16</i>	171

Part I: General Linguistics

FOLK ETYMOLOGY (IN ENGLISH AND ELSEWHERE)

Adrian Poruciu

Abstract: Folk etymology is too complex a phenomenon to be dismissed as a mere source of mistakes based on ignorance. The present paper reflects the opinions of a Romanian linguist (and Anglicist) on irregular phonetic-semantic shifts produced by inter-language borrowing, or by intra-language passage from one period to another.

A dictionary definition of folk etymology reads: "A change in form of a word or phrase, resulting from an incorrect popular notion of the origin or meaning of the term, or from the influence of familiar terms mistakenly taken to be analogous" (AHD, s.v. *folk etymology*). What we have here is mainly a negative definition. However, from a psycholinguistic standpoint, the motivation of the phenomenon may appear as positive: such formal-semantic deviations are born of the (less-cultured) speaker's desire for meaningful transparency, i.e. of a quite natural desire for control over language. Romanian peasants, for instance, have tried to make recent borrowings more comfortable: *cooperativa* 'co-operative shop' is forced into the family of the autochthonous *a cumpara* 'to buy', whence the rural *cumparativa*; the literary *neglijent* 'careless' (from French) is sometimes used as *negrijent* (as if derived from the better known *grija* 'care'); the recent *a enerva* slips into *a narava* (as if derived from *narav* 'bad habit'), and so on. It would seem that such academically undesirable shifts (of the same nature as Fr. *chaise longue* > Engl. *chaise lounge*) occur only when words are borrowed from other languages; nevertheless, folk etymologies also appear in the passage from one period to another of one and the same language. Illustrative examples are to be found especially in long-attested languages, such as English.

As is to be expected, folk-etymological shifts are more numerous in periods when (or in areas where) there are no, or very few normative factors, such as schooling and scholarly preoccupation with language. Thousands of such shifts certainly occurred in epochs we cannot check, for lack of written records; so, many of the "normal" words we use today are as they are not by law-like evolutions, but by mistakes of our illiterate ancestors. Generalized literacy will not, however, put an end to folk etymology; the above-mentioned Romanian examples, and much of the material analysed below demonstrate that the phenomenon under discussion is alive and well at the end of the 20th century.

Not only common words, but also proper names can have suffered distortions that have remained unchecked and uncorrected. Many such names are to be found in England. Beside the common

classification of English toponyms according to origins (pre-Celtic, Celtic, Roman, Anglo-Saxon, etc), we may try a general division into transparent and opaque. Folk-etymological products appear, rather paradoxically, among the transparent names, since, in their deviated forms, they obviously evoke something recognizable in the common language (though that something may be quite far from the "correct" original meaning).

What today appears as *Oxted* was earlier attested as *Oakstead* (this and most of the following toponymic examples are from the final chapter of Pinsker 1963). A certain *Smithfield* should not evoke any Anglo-Saxon smithery, since it comes from O.E. *smethe feld* 'smooth field'. And what once referred to the sea in *Mersea* is not its final part, but the initial one, since the name comes from O.E. *meres eg* 'sea's island' (i.e. genitive of *mere* 'sea' + *eg* 'island'). In such cases we become aware of diachronical twists only when there are records of older versions of one and the same place-name, or when we have sufficient clues for credible reconstructions.

Sometimes, toponymic folk-etymologies should be clarified both diachronically and interlinguistically. For instance, the name of an important Roman colony in Britain, *Eburacum*, may be interpreted as a Latinized form of a Celtic place-name derived from Celt. *eburos* 'yew' (we may also think of the continental tribe of *Eburones*; and the possibility of a pre-Celtic - "Iberian"?- source should not be excluded). The Northumbrian Angles turned the Celto-Roman name into *Eoforwic*, thus making it appear as a transparent compound in their own language: O.E. *eofor* 'wild boar'+ O.E. *wic* 'village' < Lat. *vicus* (cf. Poruciuc 1987:61); without the old attestations, today's form of the same toponym, *York*, would be totally opaque. But, we should not forget, today's clarity does not automatically mean "diachronical correctness": a place now known as *Rosedale* may never have had any roses in it, since, in the language of Scandinavian intruders in Britain, *hrosa dalr* meant 'valley of horses'.

Very interesting are the deviations caused by archaic grammatical markers which have been preserved as (unfunctional) parts of certain names. More recent speakers have tried to integrate those obscure fossils in intelligible constructions. In that vein, today's *Deanham* appears to be a clear member of the British toponymic system; but its earlier form had nothing to do with either *dean* or the toponymic suffix *-ham*: the place was called, of old, *aet denum* 'at the valleys' (the old dat.pl. marker *-um* being later "heard" as *-ham*). A place the English know today as *Bestwall* was not called like that for any special qualities of its walls; its form folk-etymologically derives from O.E. *be eastan walle* 'by the eastern wall'; so, *Best-* grew from a fusion of an old preposition and an adjective. In its turn, the semi-opaque *Attercliff* comes from an O.E. combination of preposition + definite article (dat.sg.fem.) + noun: *aet thaere cliffe* 'at the cliff'. We

may observe that place-names have been more successful than common words in preserving particles of archaic grammar (in form, though not also in meaning and function).

In the field of appellatives one can also detect folk etymologies of interlinguistic extraction, as well as of the category that may be diachronically accounted for. For instance, *buttery* 'a pantry or cellar' only seems to be related to *butter*: for a correct etymology, we should go rather to *bottle*, since *buttery* comes from O.Fr. *boteri*, derived from O.Fr. *bot* (< Lat. *buttis* 'casket'), which is also the origin of O.Fr. *botel*, wherefrom *bottle* in English. As regards internal (inter-period) shifts, a good example is today's *sandblind*, which might be misinterpreted as an echo of Sand-Man nursery traditions, or of some sand-in-your-eyes ritual; in fact, *sand-* in the (now) semantically obscure compound is a distorted development of O.E. *sam-* (a cognate of Lat. *semi-*); *sandblind* once meant simply 'half-blind'.

Many of us may remember Lewis Carroll's ever-sleepy *dormouse* (at this point I had to correct my own mis-typing: *doormouse!*). Dictionaries are quite uncertain about the name of that rodent, but there seems to exist some acceptance of a derivation from Fr. *dormir* + Engl. *mouse* (a kind of zoologically sustained etymology, since the animal is rather nocturnal and, as a hibernator, it is often found asleep during the cold season). I think that, in fact, we do not have a "real" French-English lexical contribution here, because the *-mouse* part of the compound represents rather a popular interpretation (and Anglicization) of the final part of Fr. *dormeuse* 'sleeping one'; basically, it is the situation of a French derivative (*dorm-* + suffix *-euse*) wrongly divided and made transparent by the English common-speaker's mind. Etymologically, Engl. *dormouse* may be regarded as a correlative of Rom. *dormeza* 'sofa', itself a recent borrowing of Fr. *dormeuse*. The analysis of *dormeuse* by the English mind may look strange (to linguists); or, at least, stranger than the analysis of *Hamburger (steak)*, which led to the impressively productive *burger* in American English. I dare suppose that only few contemporary Americans are aware of a link between the German city of Hamburg and their daily *cheese-, fish-, or whatever-burgers*. What I am quite certain about is that the misapprehended *ham-burger* ensured, for the second member of the supposed "transparent compound", a position in the English word-formation pool; and the process was sustained by the evocatory power of the pseudo-*ham* in the primeval *Hamburger steaks*.

There are instances of English doublets produced by folk etymology: one word may stand for the preservation, by normative factors, of a form as close to the original as possible, while the other member of the doublet (i.e. the one that looks more English) represents the popular usage. A good illustration is *asparagus/sparrow grass*: the latter came to be so widely used that it forced its way into some dictionaries. And we should not be surprised if some English

speaking wit offered even some logical explanation, such as: 'a herb beloved by sparrows'; because there are many cases of phonetic-semantic alterations suggested by actual features of the referent. For instance, I suspect that M.E. *ellebre* (< O.Fr. < Lat.-Gk. *elleborous*) became today's *hellebore* most probably because it is "hellish" (i.e. poisonous). But *shamrock* (an Irish national emblem) can certainly not be interpreted as a "sham rock"; the diminutive derivative from O.Ir. *semar* 'clover' suffered a misleading Anglicization. In other cases we may discover funny "pleonastic" developments: e.g. *chickpea* is the English name of an Asian-Mediterranean plant with pea-like seeds; in older forms of the word, it was the opening, not the final part which meant 'pea', since the M.E. version was *chiche* (< O.Fr. < Lat. *cicer* 'pea' - cf. *Cicero*!); one may, unscientifically, conclude that some Modern English chick semantically swallowed the Roman peas.

Maybe the best collection of folk etymologies could be obtained by a thorough study of overseas English. By moving to new environments, less educated English speakers were bound to get into lexical want when they came upon lots of unfamiliar things. Two main sources were at hand in such situations: (1) the former was the linguistic material the immigrants brought with them and cast into new coinages, whenever necessary (e.g. unfamiliar plants of North America were metaphorically called *black-eyed Susan*, *blue-eyed grass*, *Dutchman's breeches*, *jack-in-the-pulpit*, etc); (2) the latter source were the idioms of Red Indians, Aborigines, Maoris, or of earlier European immigrants (Spanish, French, Dutch); foreign words were often folk-etymologically deformed when Anglicized. For instance, a North American rodent is called *woodchuck*, not because it lives in the woods, but by an "adjustment" of Cree *ocek* (< proto-Algonquian **wecyeka* 'fisher'). An Australian parakeet known as *budgerigar*, colloquially *budgie*, has a similar status. Linguists may demonstrate that the bird's name comes from Aborig. *budgeri* 'good' + *gar* 'cockatoo', but, in the Australian speaker's mind, *budgie* certainly evokes *to budge* 'to move, to alter position', meanings sustained, probably, by the actual behaviour of that kind of bird; *budgie* found its subconscious correlative, just as the Oriental loanword *mohair* "naturally" found its rhyme in English, very much because it happens to designate a *hairy* material (I may go even go further, and assume the existence of a *mow-hair* form in the English-speaking subconscious). To come back to bird-name interpretation, *cockatoo* itself has only a folk-etymological relationship with Engl. *cock*, as the former derives (through Dutch) from Malay *kakatura* (cf. Dutch *kakatoe* and Germ. *Kakadu*, both opaque and closer to the original than Engl. *cockatoo*).

Sometimes folk-etymological deviations may not only be fixed in a (receiving) language, but also become international by subsequent borrowing. Old examples are *orange* and *lute*. Arab. *naranja* (< Pers. *naranga*, of Indian extraction) was better preserved by Span. *naranja*; but in Old French the word became *orange*, either because the initial

n- was misinterpreted (by French speakers) as some kind of indefinite article, or by contamination with O.Fr. *or* 'gold' (in connection with the actual color of the referent); English then borrowed the already deviated version from French. In its turn, *lute* comes from O.Fr. *leut*, a borrowing of Arab *al-'ud* (literally 'the wood', indicating the material the musical instrument was made of). The process here is different from the one which produced *orange*: the Arabic article was felt as part of the word, which may account for the perpetuation of initial *l-*. I may observe that, through another filter (Turkish or Greek), Romanian acquired *alauta/lauta* > *lautar* '(Gypsy) musician', which reflect similar adjustments of foreign lexis. A quite recent international success is that of *mangrove*: the beginning of its European career was the borrowing, by the Portuguese, of a word (transcribed as *mangue*) from a Caribbean idiom; from Portuguese, the word entered English, where it was made to remind of the familiar *grove*; then the new coinage was borrowed by French, and from there we now have it in Romanian, as *mangrove*, with the spelling pronunciation /man'grove/.

Since I have already touched the phenomenon of dictionary-fixed folk etymologies, I may add that standard English contains remarkable examples in that respect. Famous fixations were performed especially during the Johnsonian epoch, when grammarians and lexicographers became more authoritative (as standard makers) than poets and playwrights. Some Early Modern English scholarly "corrections" affected spelling only: M.E. *soverein* (cf. Milton's *sovran*) was adorned with an extra letter, becoming *sovereign*, as if related to *reign*, though the O.Fr. source-word had derived from the family of Lat. *super*, not from that of *rex* and *regnum*. By a similar cosmetic intervention, O.Fr. *amiral* (from Arab. *amir-al*, literally 'commander of') was "Latinized" as *admiral* in English, the mistaken association being with Lat. *admirari*; in this case, the graphemic bookish addition (like the "restored" *c* of *perfect* < M.E. *parfit* < O.Fr.) did not remain mute: the *d* of *admiral* is now pronounced, unlike the Latinizing *g* of *sovereign*, or the *b*'s in *debt* and *doubt* (< M.E. *dette* and *doute*, both deriving from O.Fr., not directly from Latin). What term should one apply to such deviations? "Scholarly folk etymologies" would sound rather paradoxical. It is true that, through centuries, scholars have taken pains in organizing and explaining lexical material, but examples like the ones above suggest that even the most serious dictionaries may perpetuate old errors. Psycholinguistically again, it is hard to say which is easier "to forgive": the popular integrative tendency (which tries to turn alien words into familiar ones), or the scholarly unjustified re-spellings, meant to suggest origins of prestige.

The other day I was wondering at the resemblance between Rom. *damigeana* and its English correspondent, *demijohn*. I suspected that the latter half of the English word was folk etymological.

Dictionaries (AHD, OED) inform us that the most credible source is a metaphorical *dame Jeanne* of France (so, *demijohn* should not suggest that the English are inclined to use only half a john). The *dame-Jeanne* explanation looks Roman(t)ic enough, and it becomes even more so when one finds out that Rom. *damigeana* is supposed to be a borrowing of It. *damigiana* (cf. DEX), or, according to others (cf. Sala/Vintila-Radulescu 1981;197), the Romanian word comes from Persian, through obscure intermediaries. I find that good old *bottle* is safer, to be sure.

REFERENCES

- AHD - The American Heritage Dictionary of the English Language, 1973
- DEX - Dictionarul explicativ al limbii romane, 1975, Bucuresti: Editura Academiei
- OED - The Shorter Oxford English Dictionary on Historical Principles, 1962
- Pinsker, H.E., 1963, Historische Englische Grammatik, Munchen: Hueber
- Poruciuc, A., 1987, A Concise History of the English Language, Iasi:UP
- Sala, M. & I. Vintila-Radulescu, 1981, Limbile lumii, Bucuresti:Editura Stiintifica