Winkin’, blinkin’, and nod:  
A study in historical-comparative semantics of Southeast Asian languages

Robert S. Bauer  
La Trobe University  
Melbourne, Australia

The term PANASEA (Pan-Southeast Asianism) refers to phono-semantically similar lexical items with wide geographical and linguistic distribution throughout the SEA region. Since these items must have sources, PANASEA also applies to hypothetical lexical roots which cannot be uniquely assigned to any recognized language family. Among some languages belonging to the Sinitic, Tibeto-Burman, Kadai, Austronesian, Yao, and Mon-Khmer groups, the lexemes wink (eye), wave (hand), flash (lightning) appear to constitute a word family. The PANASEA KVSJVP "to move up and down; to flash on and off" is proposed as a possible source of forms with these meanings.

1.0 Introduction

This paper again focusses attention on the phenomenon of phonosemantic similarity in the Southeast Asian linguistic area. The term refers to the similarity in phonetic shapes and meanings of two or more lexical items belonging to different languages whose genetic affiliation the linguist may or may not know. For example, I regard as phonosemantically similar the items listed below:

Language X (= Tankhur Naga, Tibeto-Burman) kʰəjəp "to wink (eye)"
Language Y (= Siamese/Bangkok Thai, Kadai) kʰajip "to wink (eye)"
Language Z (= Malay, Austronesian) kʰəjəp "to blink (eye)"

There are at least four possible explanations to account for phonosemantic similarity:

---

1 This paper was presented on October 7, 1991, at the opening plenary session of the 24th International Conference on Sino-Tibetan Languages and Linguistics at Ramkhamhaeng University, Bangkok, Thailand; a preliminary version of the paper was published in La Trobe University Working Papers in Linguistics 4 (1991):53-71. I am most grateful to my colleague, Dr. Graham Scott, for transforming my original Chaiter file into Macintosh Word format and for editing the ms. for LITUFPL 4. My special thanks to Prof. W. L. Ballard of Ehime University, Matsuyama, Japan, for making very useful comments and corrections on the working paper.
(1) *Cognition*: The lexical items are similar in sound and meaning because they are cognates, i.e., they are reflexes of the same etymon, and the languages to which they belong are genetically related.

(2) *Borrowing*: The items may be similar because one of them has been borrowed from the other language or from some third language; borrowing results from language contact and occurs independently of the genetic affiliation of languages.

(3) *Linguistic Tendency*: The items may be similar through some universal tendency, for example, sound symbolism.

(4) *Coincidence*: The items may be similar as the result of chance.²

1.1 *Cognition*

Language forms from six Southeast Asian language groups are compared in this paper; these groups include Southern Sinitic (southeastern Chinese dialects), Tibeto-Burman, Yao (= Mian), Kadai, Austro-nesian, and Mon-Khmer. These six groups in turn represent three genetic families: Sinitic and Tibeto-Burman form the two main branches of Sino-Tibetan. According to Benedict (1975), Yao, Kadai and Austronesian belong to a broader genetic grouping termed Austro-Tai. Mon-Khmer is one branch of Austroasiatic. However, Sagart (1990) has assembled a large body of lexical comparisons which he claims proves the genetic relationship of Chinese with Austronesian. Xing (1991) has examined Sagart’s material and agrees with his proposed Proto-Sino-Austronesian family. On the other hand, while acknowledging that there may have been “a very early contact relationship between [Austronesian] and Chinese”, Matisoff (1992:159-160) has raised objections (viz., insufficiently rigorous phonological and semantic correspondences and no focus on basic vocabulary) to Sagart’s methods for establishing a genetic link. In the present paper, I accept Benedict’s genetic classifications of Southeast Asian languages; however, given the long history of migration of peoples and dispersal of their languages throughout the area and the contact situations which have resulted (more on this topic in the next section), the genetic pigeonholing of languages may not contribute much to the difficult task of identifying the proto-language origin of lexical roots, when these exhibit widespread distribution across diverse groups of Southeast Asian languages.

1.2 *Borrowing*

At almost any time in their histories, languages are in contact with other languages. This is the natural state of affairs. Contact borrowing is a pervasive, continuous, natural process that causes languages to change, whether it be to acquire a new vocabulary while adopting a new cultural area, such as sport, religion, technology; to develop or lose a tone system; to switch

² For a similar list of “four possible explanations”, see Matisoff 1976:265-8.
from SOV to SVO word order. Any area of language can be affected by borrowing and the influence that languages in contact exert on one another pays no regard to their genetic affiliation. A common characteristic of linguistic areas around the world is the phenomenon of convergence: as the result of contact-borrowing, typologically distinct and distantly related languages become more like their neighbors in vocabulary, phonetics, or grammar. The Southeast Asian linguistic area is particularly noted for convergence and the concomitant problems that convergence creates for the historical linguist. Attempting to separate out native words from borrowings in order to identify the genetic affiliations of languages or to assign lexical roots to particular proto-languages presents the historical linguist with formidable difficulties in “the hothouse homogenizing atmosphere of Southeast Asia” (Matsoff 1983:63) where the migratory movements of peoples over the millennia have created myriad opportunities for convergence through contact borrowing among the area’s many languages. (In this regard, it is worth noting that in Trubetzkoy’s view [summarized in Renfrew 1987:108] areal convergence in Europe was sufficient to explain the similarity of European languages and thus the reconstruction of a Proto-Indo-European ancestral language was made unnecessary).

At any rate, the notion that so-called core or basic vocabulary is not subject to replacement through borrowing has become the cardinal principle to which Proto-Worlders, Nostraticists, and others have anchored their reconstructions of remote linguistic relationships. Guided by the claims of Dolgopolsky (1986) in their reconstruction of Nostratic, Kaiser and Shevoroshkin (1988:311) have explicitly stated this “principle” as follows:

Based on the material of more than 250 languages, Dolgopolsky’s study shows that there is a hierarchy of stability of lexemes—i.e., certain lexemes are more resistant than others to replacement by other lexemes (in the same language) and to borrowing into other languages. When two languages come into contact, certain words are easily borrowed (cultural items), while others are seldom, if ever, subject to borrowing. Dolgopolsky composed a list of fifteen most stable lexemes, which we repeat here: 1. “I, me”; 2. “two, pair”; 3. “thou, thee”; 4. “who, what”; 5. “tongue”; 6. “name”; 7. “eye”; 8. “heart”; 9. “tooth”; 10. verbal NEG (negation and prohibition); 11. “finger/toenail”; 12. “louse”; 13. “tear” (n.); 14. “water”; 15. “dead”.

I readily acknowledge that one can semantically distinguish between basic vocabulary, on the one hand, i.e., lexical items which every language will almost always possess, such as bodypart names, deictics and kinship terms, and non-basic vocabulary, on the other, i.e., names for cultural items, e.g. soap, chocolate, which are easily transmittable via the process of contact borrowing. Just how useful is this distinction? Consider verbs associated with body parts or natural phenomena, such as wink and lightning, which this
paper is concerned with. In my view these will also need to be classified as core vocabulary—they would seem to have more in common with words belonging to this category than non-core, culture words such as soap and chocolate which have both achieved global distribution (Bauer 1992a). In the final analysis, the claim that core vocabulary is less likely to be replaced by loanwords remains so much wishful thinking on the part of those linguists who gingerly clutch this convenient, heuristic compass to guide them through the convoluted pathways of the historical comparative jungle. As messy as it makes things, we must accept that any lexeme is grist for the borrowing mill—even bodypart terms.3

1.3 Linguistic Tendency

By this general phrase I mean sound symbolism. How much this contributes to my topic I cannot say here. In an earlier paper (Bauer 1988b) I proposed that the tendency of such lexemes as mother, father, breast, milk, and suck (which apparently constitute a word family) to show megalophonosemantic similarity across many languages of the world4 ultimately derives from shared patterns of sound symbolism associated with infants sucking their mothers’ breasts (curiously, Dolgopolsky (1986:29) omitted linguistic tendency from his own list of three explanations for phonosemantic similarity). Furthermore, given a particular lexeme from this set of five (which can probably be expanded), one can make predictions about its phonetic shape and the possible phonetic contrasts in initial consonants that the other lexical members of this set may display in various language groups.

Shared patterns of sound symbolism for another group of lexemes meaning to close the eye, to wink the eye, to flash lightning may partially explain both the origin of these words and the phonosemantic similarity they exhibit across diverse groups of Southeast Asian languages; that is to say, different groups of speakers may be using similar speech sounds to symbolize the rapid movement of blinking eyelids or the sudden flash of lightning in the sky. Speakers may even perceive some direct iconic relationship between the closure of the eye and the articulatory gesture of closing the mouth to produce the final bilabial stop -p which is shared by (almost) all the lexical forms cited in this paper.5 To the extent that sound symbolism lies below the level of conscious

---
4 Bauer 1992b:147 illustrates the phenomenon with forms for “breast” from 18 languages representing eight language families.
5 During the 1991 Sino-Tibetan Conference James R. Chamberlain kindly gave me a most intriguing paper he had written, entitled “Frog mouths and mimesis: an essay on the relationship between form and meaning” (no date). In this paper he demonstrated what he believes to be the non-arbitrary link between meaning and sound for a large group of Lao words ending in -p and having to do with the meaning “closure”. To me what he is proposing is a Lao phonetheme -p “to close; closure”. In reading through his paper, I was struck by the similarity of forms from southern Chinese dialects to the Lao forms. It is quite possible that similar