A Note on Eliciting Words

Karl J. Franklin
GIAL and SIL International

Many years ago Arthur Capell (1952) wrote a small booklet which was useful in eliciting words and sentences. Although I have since lost the book, it was useful for eliciting vocabulary for the first time in an unwritten language. As a field linguist I soon learned that there are many pitfalls in eliciting even "basic vocabulary."

Bee and Pence (1962) first tried to standardize a word list that SIL field workers should use in their first attempts to transcribe an unwritten language. The idea was that the standard list could then be used for comparative and lexicostatistical purposes. They also constructed a "supplementary" list of items that were not in the normal SIL list, but were often found as basic vocabulary, especially in the lowlands of Papua New Guinea.

Later Don Laycock (1970) elaborated on the difficulties of eliciting materials in even such an assumed standard vocabulary list. He examined 357 entries, conflating those from two Swadesh lists (1952, 1955), Wurm's New Guinea Non-Cultural Vocabulary[1], the standard SIL list, the list used by McElhanon (1967, based on the SIL list), the Linguistic Questionnaire of the Tri-Institutional Pacific Program (Department of Anthropology, Yale University, 1954[2]), and Laycock's own list, selected from Wurm's materials.

Of the 357 items that Laycock examined, combining all of the word lists given above, he judged only 56 as highly reliable and stable, but decided that another 23 should always be elicited. By reliable he meant that the word presented little difficulty in elicitation, and by stable he meant that the item tended to show cognates throughout a large number of languages that were found to be related. He concluded that 17 items were highly unsuitable for elicitation (i.e. difficult to be sure of the meaning) and that "less strong objections" could be raised against another 137. These also led to questionable meanings in a number of languages. He simply discarded the remainder and concluded that only 233 words of the 357 he examined were the best to be included in word lists for elicitation.

His parting words were that "it is no easy matter to devise a list of even 100 words for lexicostatistical classification in New Guinea, and that little faith should be put in detailed linguistic subgrouping arrived at by this [lexicostatistical] method" (Laycock 1970:1173).

Laycock (following McElhanon 1967) omitted words that linguists had found to often involve repetitions of the same vernacular term, such as dirty-black, far-long, near-short, feather-hair, sharp-tooth, narrow-thin-little, wide-thick-big, fog-cloud, narrow-thin, wide-thick-big, river-water, here-this, there-that, wife-woman, husband-man, lie-sleep, wipe-wash, hear-know, and kill-hit. To this could be added meat-flesh, blood-red, drink-eat, bark-skin, river-water, stone-mountain, egg-grease, sun-moon, tree-fire, yesterday-tomorrow, seed-eye-flower, person-man, finger-thumb, and probably many more. However, not all Papuan languages distinguish semantic pairs or triplets in the same way, underscoring the fact that a word is a mental image, based on cultural criteria, so there is unlikely to be uniformity except in a small number of cases. Often additional knowledge of the language and semantic probing

---

[1]Never published by Wurm, but published, with comments, by Laycock in 1965.
[2]According to Laycock (1970:1128) this is Section 1: Basic Vocabulary, 215 items, of the unpublished questionnaire.

will distinguish many of the semantic pairs.[3] And, of course, not all languages in the Highlands of PNG have the same list of repetitions that Laycock (or McElhanon) list.

My purpose here is not to debate these assessments, or the problem of "multiple cognates (Wurm and Laycock 1961/62:134, discussed in McElhanon 1967:126-27) but rather to point out how difficult it is to elicit anything in a language that one is recording for the first time. I only wish to emphasize that words represent cultural objects, events, themes, and so on, depending on the perspective of the native speaker and hearer, so we can expect different words in the same apparent situation.

It follows that different researchers who elicit the same lists in the same language may also vary considerably. For example, Reesink (1976) and I (1973) differed considerably (as much as 25%) in the words in our lists for languages along the Aramia River of the Gulf Province, and this led to our calculating different degrees of relationships for these languages.

Word lists are necessary because they help us when there is a problem hearing longer stretches of speech, especially if someone is hearing a language for the first time. However, compounds are difficult to recognize because they require some grammatical analysis as well. For example, in my own early field notes in Kewa, I find many instances of what we now recognize as lexical compounds that we had written as either one word or two. The following chart gives some examples, with comments on the compounds.

Table 1: Compounds in Kewa elicited as one word or two

Written as one word	**Written as two words**
toiri (*to* + *iri*, 'body' + 'hair' = 'hair, not on the head'). This is a compound, *to-iri*, because 'body' does not occur with the possessive marker.	*aa kidipa* ('foot' + 'nail' = 'toenail'). This again is a compound, *aa-kidipa*, and parallels *ki-kidipa* ('hand' + 'nail' = 'fingernail').
kibuni (*kibu* + *uni*, 'shin' + 'bone' = 'shinbone'). Heard as one word because the first word ends in /u/ and the second begins with /u/. Better written as a compound, *kibu-uni*.	*aariyapa* (*aari* + *yapa*, 'mountain' + 'daylight' = 'lightning'). A compound, *aari-yapa*, but not initially recognized because of /y/ between the vowels /i/ and /a/.
ekepo (*ee* + *kepo*, 'old garden' + 'fence' = 'garden fence'). Vowels that occur by themselves as words are lengthened, so *ee-kepo* would be the correct transcription. The form for 'old garden' is also not marked as a possessive phrase.	*ipa ini* (*ipa* + *ini*, 'water' + 'eye' = 'pond', 'lake'). Better written as *ipa-ini* to show that it is a single lexical item meaning simply 'lake' and not a possessed form meaning 'the eye of the water'.

[3]For example, in Kewa one should note the differences between the flesh of animals and humans, the fact that adjectives modify some particular object, that rivers are named, that wife and husband are distinguished by kinship terms, that wash may be a loan word, that stone and mountain depend on terrain, and so on. Any word spoken in isolation has to be defined on the basis of its semantic context. For example, it is commonly assumed that tree and fire are semantic synonyms in many Highland languages; in Kewa (and other Engan languages) the same form (*repona*) will be elicited for both. However, once we elicit 'coals', 'embers', 'ashes', and kinds of common trees or kinds of firewood, it soon becomes obvious that there is a great range of words for both "tree" and "fire."

yaaga (*yaa* + *agaa*, 'bird' + 'mouth' = 'beak'). A compound, *yaa-agaa*, that contrasts with *yaga* 'chin'.	*yaa ada* (*yaa* + *ada*, 'bird' + 'house' = 'bird's nest'). A compound, *yaa-ada*, that contrasts with *yaa-na ada* 'the house of the bird' or 'bird house'.
etane (*etaa* + *ne*, 'food '+ 'sharp' = 'arrow'). A compound but not recognized initially. Now written as *etaa-ne*.	*pepena yae* ('decorate' + 'something' = 'ornament'). A compound, *pepena-yae*. 'Something decorated' would be *yae-na papena*.
yapaapu (*yapa* + *aapu*, 'possum' + 'tail' = 'headband'). A compound, *yapa-aapu*, and not *yapa-na aapu* 'the tail of the marsupial'.	*mena ami* ('pig' + 'teat' = 'fat for greasing body'). Written as the compound *mena-ami*; 'pig fat' is *mena-na ami*.
katipi (*kati* + *pia*, 'gray colored clay' + 'sitting' = 'gray hair'). Now written as the compound *kati-pi*.	*nopare* ('down' + 'location' = 'low'). Written as either a compound or a deictic + a locational clitic.
seketa (*sete* + *la*, 'wet' + 'saying' = 'wet'). Written as *seke-ta*, because, if written alone, would mean 'he is saying it'.	*ki mea* (*ki* + *mea*, 'hand'+ 'get' = 'shake hands'). Written as a clitic: *ki-mea*. Separately the words would mean 'fetch the hand'.

Some of the words that Laycock, Wurm, McElhanon and others have suggested as problems may turn out to show relationships that are not apparent at first. For example, Dutton (1973) collected and compared the "cultural" vocabulary items, including common foodstuffs, such as sweet potato, taro, yam, banana, sugarcane and associated vocabulary, such as garden and fence, and arrived at the conclusion (1973:494) "that all except those individual forms or small sets of related ones which are restricted to single language or members of language families or neighbouring languages… are borrowed and are therefore justifiably regarded as 'cultural' …."

His study showed clearly the pervasive and extensive borrowed nature of these terms and the various phonetic and lexical changes that take place in the process.

A final point: although eliciting word lists is necessary in early fieldwork, it often obscures the many semantic relations and new terms that occur and that we will find with further study. For example, the word for 'moon' is different in the three dialects of Kewa: *eke* in East Kewa, *akua* in West Kewa and (often) *suba* in South Kewa and other areas. We might conclude that only the form we elicit is used in a particular dialect. However, although *akua* is 'moon' in West Kewa, a 'full moon' is *eke kura* and the first quarter of a mood is *suba maari*. In other words, the three dialect free forms are all used in descriptions of the moon in its various stages.

References

Bee, Darlene and Alan Pence. 1962. Towards standardization of a survey list for Papua and New Guinea. *Oceania Linguistic Monographs* 6:64-75.

Capell, Arthur. 1952. *Methods and materials for recording Papuan and New Guinea languages*. New Guinea Department of Education, Official Research Publication No. 1. Sydney: Government Printer.

Dutton, Tom E. 1973. "Cultural" items of basic vocabulary in the Gulf and other districts of Papua (Part 1: Foodstuffs and associated agricultural terms). In Franklin, ed., pp. 411-538.

Franklin, Karl J., ed. 1973. *The linguistic situation in the Gulf District and adjacent areas, Papua New Guinea*. Pacific Linguistics C-26. Canberra: Australian National University.

Laycock, Donald C. 1965. *The Ndu Language Family (Sepik District, New Guinea).* Pacific Linguistics C, No. 1. Canberra: Australian National University.
Laycock, Donald C. 1970. Eliciting basic vocabulary in New Guinea. In S. A. Wurm and D. C. Laycock, eds., *Pacific linguistic studies in honour of Arthur Capell*. Pacific Linguistics C-13:1127-76. Canberra: Australian National University.
McElhanon, Kenneth A. 1967. Preliminary observations on Huon Peninsula languages. *Oceanic Linguistics* 6:1-45.
Reesink, Ger P. 1976. Languages of the Aramia River area. Pacific Linguistics, A-45:1-37. Canberra: Australian National University.
Swadesh, Morris. 1952. Lexicostatistic dating of prehistoric ethnic contacts. *Proceedings of the American Philosophical Society* 96:452-63.
Swadesh, Morris. 1955. Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics* 21:121-37.
Wurm, Stephen A. and Donald C. Laycock. 1961/62. The question of language and dialect in New Guinea. *Oceania* 32:128-43.