The data and hypotheses presented here are part of a broader investigation concerned with numeral classifier systems considered both as representatives of a type of nominal classification and in relation to the problems of quantification in language, what is meant by quantification in this connection is the manner in which languages express the fact that reference is being made to a quantitatively delimited amount of the thing mentioned.

Such a typological approach involves both synchronic and diachronic considerations. Initially, we take into account an extensive, ideally, an exhaustive sample of languages which is based on preliminary notions regarding the definitional characteristics of the type. A comparison of such languages leads to a number of synchronic generalizations, usually implicational in form. The second major aim is to uncover the dynamic principles, that is the recurrent types of change in historically independent instances involved in the rise, subsequent expansion and ultimate dissolution of the type. In carrying out this part of the investigation our methods include deductions based on internal reconstruction within individual languages, the comparative method within linguistic stocks and direct historical documentation where this is available.

As noted initially, the tentative conclusions presented here are but a portion of a broader study which is in progress. In the present study the emphasis will be on questions relating to the initial conditions under

---

2 The present research was supported by the National Science Foundation as part of the Language Universals Project at Stanford University.  
3 The somewhat vague term ‘thing mentioned’ is used here because although the present study is basically confined to nominal phrases, verbal action can also be quantified. This is briefly discussed in the final section of the paper.  
4 For discussion and exemplification of these methods, see Greenberg 1966, 1969, 1970(a), 1970(b).
which numeral classifier systems may be conjectured to arise. In the final section, in order to place the present study within the more general perspective of the study as a whole, a series of other problems and in some cases hypotheses regarding them will be outlined without pursuing them in detail.

One limitation should be mentioned at the outset. Systems of the type with which we are concerned here have undoubtedly, in some instances, arisen under conditions of language contact. For example, in those Dravidian languages which have such systems it seems clear that they have developed in general as a result of contact with Indo-Aryan languages. Any theory of origin will ultimately have to take into account both the conditions under which pristine systems arise and those in which contact is a major factor. In the study in the present form, the predominant emphasis is on the former.

As mentioned earlier the sample is not exhaustive and this, of course, adds still further to the tentative nature of the results. Nevertheless the hypotheses presented here are based on quite extensive data. They are presented, here, in the hope that they may provide at least a basis for conclusions that can be tested and modified in the light of both raw data and more penetrating theoretical analysis.

We begin with an attempt at a preliminary definition of what constitutes a numeral classifier language in terms of the existence of a particular syntactic construction. A considerable number of the world’s languages including almost all of these in Southeast Asia exhibit the following characteristic. An English phrase such as ‘five books’ is rendered in translation by a phrase containing, outside of possible grammatical markers, not two but three elements. The kind of literal translation often supplied

---

5 On this topic see especially Emeneau (1956).
6 A list of languages in my sample follows. In a few instances the numeral classifier system is very marginal, e.g. Bulgarian, because of the use of duši ‘soul’ used in enumerating persons, and Hungarian, because it has a numeral series used only with persons: Ahom, Ainu, Assamese, Banggais, Bengali, Black Thai, Bodo, Breton (Medieval), Bribri, Brou, Bulgarian, Burmese, Cebuano, Chinese (Archaic, Mandarin, Hakka, Cantonese), Chiripo, Cholon, Chontal (Mayan), Cuna, Day, Diao, Engeni, Egyptian Arabic, Empeo, Fijian, Garo, Gilbertese, Gilyak, Guaymi, Haida, Hausa, Hungarian, Hupa, Iban, Ibibio, Irish, Ishkashim, Jacaltec, Japanese, Kachin, Karen, Karo-Batak, Katu, Kei, Khamti, Kharit, Khasi, Khmer, Khmu, Kiriwina, Kolami, Korean, Kurukh, Laotian, Lisu, Malay, Man, Maru, Merir, Mikir, Miri, Mon, Mota, Muchik, Nahuatl (Classical and Tetelcingo), Nauru, Ojibwa, Omani Arabic, Ossete, Palau, Palaung, Parji, Pashto, Persian, Pocomchi, Ponahe, Pur, Rawang, Samoan, Shan, Sosorol, Tajik, Taron, Tarascan, Tat, Thai, Tho, Tlingit, Toba-Batak, Totonac, Trukese, Tsimshian, Turkish, Tzeltal, Tzotzil, Uvea, Uzbek, Vietnamese, White Thai, Wolio, Yurok.