The Economics of Artificial Languages: Thoughts on the Problem of Cost Minimization

Jonathan Pool

1. Artificial Codes, Artificial Languages, and Natural Languages

The term "artificial language" is used in several senses, of which there are two major ones in scientific contexts. First, it can mean what I shall call an "artificial code": a set of explicit rules for communication which have come into existence by being prescribed. Such codes have been devised for communication among humans, between humans and animals, between humans and machines, between animals and machines, among animals and among machines. The secret codes invented by pickpockets or baseball players, and the public codes embodied in the colors of armed forces uniforms, are examples of artificial codes for human-human communication. Chimpanzees have been taught artificial codes (Rumbaugh, 1977). And hundreds of programming languages have been invented for human-machine communication (Feldman, 1979).

As used below, "artificial language" will have a second sense, differing from "artificial code". A tentative definition is: a partially explicit set of rules for communication among humans, which has evolved from an artificial code, which is usable both orally and in writing, and which allows itself to be modified to communicate any ideas that can be communicated with any natural language. In the history of the world approximately one thousand artificial codes are known to have been devised with the expressed hope that they would evolve into artificial languages (Dulič̆enko, n.d.). A few of them have actually done so to a noticeable degree; these include Volapük, Esperanto, and Ido. Today Esperanto, in fact, ranks approximately 65th among the world's languages in its rate of book publication ("Laste aperis", 1979; United Nations, 1979: 943–945).

The distinction between artificial codes and artificial languages follows Cherry's (1966: 8) distinction between sign systems and languages. Artificial languages, as defined above, resemble one of seven types of languages (standard, classical, artificial, vernacular, dialect, creole, and pidgin) that have been identified by Stewart (1968) on the basis of the presence of absence of four characteristics: standardization, autonomy, historicity, and vitality. In particular, artificial languages, according to Stewart, have standardization and autonomy but lack historicity and vitality. In other words,

they are (in part) formally codified and their evolution is not dominated by any other language, but they are not perceived as resulting from a long period of development and they are not primarily transmitted over time as native languages. However, if an artificial language were to obtain historicity and vitality after some time, it would no longer be artificial in Stewart's typology, while it would still be artificial according to my definition. Except for this difference, the term "natural language" used above is intended to correspond to any of Stewart's other six types.

2. The Economic Importance of Artificial Languages

Both artificial codes and artificial languages have important economic aspects. Among the former, it is well known that programming languages differ in the efficiency with which they can obtain certain responses from certain computers. The optimal choice of programming language requires taking into account the costs of language learning, programming, program modification, computer-language translation, computer execution, program storage, data storage, and conversion of programs across time- and machine-specific versions of the language or across languages (cf. Pratt, 1975: 6–10).

It is less known how artificial languages, such as Esperanto and Volapük, differ in their economically relevant characteristics, but a number of studies have found major differences in such characteristics between an artificial language and a natural one (e.g. French). In spite of this, there is little theory about the economics of human languages in general, and almost none about the economics of artificial languages. This is surprising when one considers that far more resources are certainly spent on the learning and use of human languages (exempting native languages) than on the learning and use of programming languages.

The paucity of glottoeconomic theory has the result that we know little about how to analyze the decisions that political authorities and ordinary individuals make about language: decisions about which languages to learn, which languages to teach, which languages to officialize, which language reforms to carry out, etc. Decisions like these have economic consequences for multilingual countries (India, USSR, Canada, etc.), for countries that use writing systems of questionable efficiency (Chinese, Japanese, English, Arabic, etc.), and for organizations whose members are highly diverse in their linguistic repertoires (United Nations, European Communities, Organization of African Unity, Union of International Associations, World Council of Churches, etc.). Such decisions impact not only total expenditures and total benefits, but also the way costs and benefits are distributed.

The absolute and relative welfare of entire countries, social classes, professions, and ethnic groups is affected, and the effects may be large (see, e.g., Fishman, Ferguson, and Das Gupta, 1968; Frank, 1971; Rubin and Jernudd, 1971; United Nations, 1977).

3. Alleged Economic Advantages and Disadvantages of Artificial Languages

It may be possible for a person, organization, set of persons, or set of organizations to obtain economic benefits of various kinds by choosing an artificial language rather than a natural language when making certain language decisions. Let us survey the types of benefits that have been alleged to exist.

a. The learning of an artificial language is less costly than the learning of a natural language. A number of studies in various countries have evaluated this claim. Although precision in the comparison of learnabilities is difficult to obtain, although most of the studies until now have been weak in the degree of experimental control achieved, and although all studies conducted so far of which I am aware have used one and the same artificial language (Esperanto), still the results have been reasonably consistent in supporting the claim. Reported and estimated ratios of learning times for the achievement of similiar levels of (active and passive) competence in Esperanto vs. natural languages have ranged from 1:2.7 up to 1:15 (Frank, 1977; Janton, 1973: 118; Kalckhoff, 1978; Markarian, 1964: 6; Rakuša, 1970: 38–39). When the research literature is combined with statements made by speakers of Esperanto whose native language are non-Indo-European (e.g. *El Popola Ĉinio*, 1980), it appears that the ratio of learning times for Esperanto vs. natural Indo-European languages does not differ substantially between native speakers of Indo-European languages and native speakers of other languages. Because of the dominant share that Indo-European languages have in the second-language market, little is known about the relative learnability of Esperanto vs. non-Indo-European languages as second languages. Wells (1978) has argued that Esperanto is an Indo-European language only in its lexicon, and that its phonology, morphology, syntax, and semantics contain attributes resembling several other language families. Whatever lexical bias toward the Romance and Germanic languages exists in Esperanto or any other artificial language may be substantially smaller than the corresponding bias of a natural language because of the high productivity of the affixes and grammatical morphemes and the wide international currency of the lexical morphemes of the artificial language. Thus undocumented claims that, for example, Esperanto cannot be learned any faster than English by a native speaker of Fijian (e.g. Farb, 1977: ch. 16) are clearly premature. Further

research, in which cross-cultural and cross-lingual comparisons are made among learners, in which prestige differences between the target languages (presumably unfavorable to the artificial one) are reduced, in which experimenter biases (presumably favorable to the artificial language) are reduced or balanced, and in which different modes and degrees of competence are examined, will help resolve this question.

b. Translation and interpretation from or into an artificial language is less costly than from or into a natural language. There are three major reasons for expecting that this claim may be true. First, if an artificial language is less costly to learn than a natural language, and if the cost of translation and interpretation reflects the cost of training required to create the manpower that performs this service, then the cost of translating or interpreting *from* an artificial language should be lower than the cost of translating or interpreting from a natural language, since in general translators and interpreters work from a second language into their native one. Second, considerable (sometimes nearly 100%) waste of translating and interpreting manpower has been reported in international organizations as a result of the unpredictability of the demand for translating and interpreting services, the rapidity with which the value of their products declines with time, the high cost of transportation among and temporary lodging at the sites of meetings, the fine linguistic specialization of existing translators and interpreters, and the failure of simultaneous interpretation (especially when conducted through an intermediate language) to convey more than a fraction (sometimes estimated at 50%) of the meaning into the target language (Piron and Tonkin, 1979; United Nations, 1977). If an artificial language can be learned well enough to be used as a target language as well as a source language, then the choice of an artificial language as target language does not constrain the choice of translators or interpreters to those having a particular native language, and the waste due to the "lumpiness" of this service, unpredictability, poor quality, etc. can be reduced. Third, it can be expected that an increasing proportion of the cost of translation in the future will be attributable to computer-initiated or computerassisted translation. The difficulty and hence cost of most aspects of the computer processing of Esperanto has been reported to be less than that of computer processing of natural languages (Maas, 1975). This is because the artificial language, being syntactically, morphologically, and semantically more regular and less ambiguous, permits analysis with smaller programs, which therefore cost less to develop, store, and execute, and which result in target-language texts that require less human correction to render them acceptable.

c. The learning of a natural language costs less when it begins with the learning of an artificial language than when it begins without such a preparatory phase. According to this claim, the learning of an artificial language

causes learners to increase their aptitude and desire for natural-language learning. Under some conditions (e.g. 500 or more hours of study of a natural language) the investment in the artificial-language preparatory phase speeds the subsequent rate of natural-language learning enough to more than repay the time invested in the preparatory phase (Frank, 1978; Frank, Geisler, and Meder, 1979; Geisler, 1980; Markarian, 1964). Hence two languages – a natural and an artificial – can be learned at a lower total cost than a natural language alone. Total cost must, of course, also include the costs of teacher preparation, teaching materials, etc.; but Lobin (1978) estimates these at only about $ 1 per pupil per year more for a two-language program than for a one-language program because of the extremely short time (1 week) required to retrain foreign language teachers to teach an additional course in an artificial language. Reasons for the impact of an artificial language on the subsequent learning of a natural language may include (a) an increased understanding of grammatical concepts resulting from the learning of a language which has a high degree of one-to-one correspondence between grammatical morphemes and grammatical concepts, (b) an increased motivation for second-language learning resulting from an initial experience of success with an easily learned language, (c) an increased proclivity to learn the most useful lexical items as a result of learning a language a high proportion of whose lexicon is of this kind (Chaves, 1979), or (d) the increased aptitude that the learning of any second language confers. If the last reason is the predominating mechanism, then any other language would do as well as an artificial one, unless the determining variable is the amount of second-language competence rather than the amount of second-language study. Research comparing an artificial and a natural language (i.e. a different natural language from the target one) as linguistic propaedeutics has apparently not yet been conducted.

d. The choice of an artificial rather than a natural language as the medium of communication in situations where there is only one such medium and in which competence in the medium is a prerequisite for engaging in a certain kind of production results in a more efficient allocation of manpower among productive activities. The argument is that there are many persons with high aptitudes in certain fields (e.g. natural sciences and engineering) but low aptitudes in second-language learning, since these two kinds of aptitudes are not strongly correlated. When the language of transnational communication in a field of production is a natural one, a larger number of persons who are fitted by aptitude for that field are excluded from it (or from prominence in it) by failing to achieve competence in that language than is the case when the language is artificial and hence not as difficult to learn. I have not discovered any statistics on the rate of disqualification for professional study on the basic of failure to satisfy foreign

or national language requirements, but this appears to be a major phenomenon in countries where a non-native language is the major medium of higher education (e.g. India) or is used as a higher education selection criterion even though not used as a medium of instruction at that level (e.g. Turkey). Noss (1967) has dealt with these problems in the countries of Southeast Asia. On the other hand, selection of an artificial language for professional communication which has previously been served by a natural language entails, as does any change of language, a period of bilingual transition, in which it is even more costly than before for the linguistic needs of the profession to be met (by language learning or translation of literature). How costly this transition period would be would depend on the rate of obsolescence of the literature and of the manpower in the field.

e. The choice of an artificial language as one medium of communication in situations where those competent in any of the permitted media of communication enjoy substantial advantages creates a political situation in which a smaller number of languages can be chosen, a smaller amount of translation and interpretation is required, a smaller total expenditure is required, a more equal sharing of the total expenditure is obtained, a larger total benefit from communication is obtained, and a more equal sharing of that benefit is achieved. These advantages of an artificial language are claimed to result from what is called its "neutrality". The argument is generally applied to the choice of official languages in international organizations some of whose members are native speakers of the current or prospective official languages. In that case, any one official language would advantage the member that speaks it natively, and the prospect of such an advantage induces other members to oppose the officialization of that language without their own language also being officialized. Thus the inexorable increase in the number of official languages that has been observed in the United Nations system (United Nations, 1977) emerges. The officialization of an artificial language would not lead to such a process, since no member would be substantially and obviously advantaged thereby. This argument, however, would not apply under certain conditions: (1) where the artificial language has obtained large numbers of native speakers and they are unevenly distributed among the members of the organization; (2) where at least one candidate for officialization among the natural languages has very few or no native speakers among the members (e.g. a multilingual state with a history of a foreign official language); (3) where the members are willing to accept the principle of non-linguistic (e.g. monetary) compensation for linguistic disadvantages.

Five alleged economic advantages of artificial over natural languages, and qualifications of them, have been summarized above. There are others in the economic realm, such as the lower cost of programming a speech syn-

thesizer to convert a standard written text into intelligible speech (Sherwood, 1979). There have also been, of course, assertions of psychologically, socially, and politically relevant differences, which will not be considered here. In spite of the primitive state of current knowledge, it appears that artificial languages have been and therefore can be devised in such a way that they can be learned, or translated from, at substantially less than the cost incurred for learning or translating from a natural language. A difference of such magnitude deserves efforts at empirical verification and explanation and, in the meantime, some analysis of how language decisions should rationally be made if two kinds of languages really exist with such major differences.

4. The Problem of Language Choice

Let us now look at some implications of different learning costs between artificial and natural languages. What follows will be only a summary of results from computations (Pool, 1980a) based on a simple model. But they will give a flavor of the kinds of results that mathematical models of linguistic economics can provide.

Suppose that the members of a group of people have existing linguistic repertoires, and that we know how much it would cost each person to learn each of the relevant languages that he does not already know. Can we determine the cheapest method of bringing the members of this group into a state of mutual "communicability"? As long as we can compare one person's costs with another's, it is possible to find the cheapest solution to the group's language barrier.

Mathematical analysis of models with such assumptions produces some non-obvious conclusions. First, the cost-minimizing solution may involve more than one language, even though we usually think of costs rising as the number of languages in use rises. Second, even if one language (e.g. an artificial one) is much less expensive for everyone to learn than another (e.g. natural) language, it *still* may be less expensive for the whole group to adopt the more expensive language; this can happen when those who already know the chosen language are numerous enough or when they experience high enough learning costs in comparison with others. A third conclusion is that inventors of artificial languages, under certain conditions, can actually reduce the relative efficiency of their language by making it similar to the natural languages that the potential users already know.

Typically, analysis shows that an artificial language becomes either part of, or the only language in, the least expensive solution to a language barrier

when it reaches a certain level of learning ease relative to the other competing languages, and when it is within a certain linguistic distance of the languages already known by the would-be communicators. What this does not tell us, however, is whether the cheapest solution to the problem will actually be adopted. This depends on how the costs of that solution are distributed, as well as whether the people involved need to coordinate their learning plans to avoid unintended conflicts. Analyses of these problems, which have only recently been started (Pool, 1978, 1980a, 1980b, 1980c), already make it clear that the problem of finding the economically optimal solution to a given language problem is much more complicated than is commonly thought, and that considerations of strategy and equity will interact in these solutions in ways that we are a long way from fully understanding.

Author Note

I am grateful for research support from the Program in Comparative Studies in Ethnicity and Nationality, the Graduate School, and the Graduate School Research Fund of the University of Washington, and from the Institut für Kybernetische Pädagogik of the Forschungs- und Entwicklungszentrum für objektivierte Lehr- und Lernverfahren, for facilitating this work, as well as for helpful references to previous work from Prof. Dr. Helmar Frank.

References

Chaves, Sylla M. (1979). Überlegungen zur Lernerleichterung im Fremdsprachenunterricht durch Vorausstellung der Internacia Lingvo. *Grundlagenstudien aus Kybernetik und Geisteswissenschaft*, 20, 122–130.

Cherry, Colin (1966). *On Human Communication*, 2nd edn. Cambridge, Mass.: M.I.T. Press.

Dulič̣enko, Aleksandr (n.d.). Bibliografija mezhdunarodnykh jazykov. Unpub. ms. *El Popola Cinio* (1980).

Farb, Peter (1977). *Word Play*. New York: Bantam.

Feldman, Jerome A. (1979). Programming Languages. *Scientific American*, 241: 6 (Dec.), 60–93.

Fishman, Joshua A.; Ferguson, Charles A.; and Das Gupta, Jyotirindra; eds. (1968). *Language Problems of Developing Nations*. New York: Wiley.

Frank, Helmar (1971). Die Sprachbarriere zwischen den Wissenschaftlern. *Umschau in Wissenschaft und Technik*, Heft 7, 236–238.

Frank, Helmar (1977). Die Meßmöglichkeit der relativen Schwierigkeit von Fremdsprachen. In *Kybernetik und Bildung III*, ed. Günter Lobin and W.D. Ekkehard Bink (Paderborn: Schöningh/Hannover: Schroedel), 14–25.

Frank, Helmar (1978). Grundlagen und sprachpädagogische Anwendung einer informationstheoretischen Transferanalyse. *Grundlagenstudien aus Kybernetik und Geisteswissenschaft,* 19, 75–88.

Frank, Helmar; Geisler, Evelyn; and Meder, Brigitte S. (1979). Nachweise des strukturbedingten Transfers aus dem Sprachorientierungsunterricht. *Grundlagenstudien aus Kybernetik und Geisteswissenschaft,* 20, 14–28.

Geisler, Evelyn (1980). Mezurado de la lernplifaciligo de la angla pro ILo. *Eŭropa Dokumentaro,* 25, 4.

Janton, Pierre (1973). *L'Espéranto.* Paris: Presses universitaires de France.

Kalckhoff, Gerhard (1978). Eine vergleichende Darstellung der Erlernbarkeit von Sprachen. *Grundlagenstudien aus Kybernetik und Geisteswissenschaft,* 19, 55–60.

Laste aperis (1979). *Esperanto,* 72, *passim.*

Lobin, Günter (1978). Bildungsökonomische Analyse zu verschiedenen Modellen des Früh-Fremdsprachenunterrichts. *Grundlagenstudien aus Kybernetik und Geisteswissenschaft,* 19, 126–132.

Maas, Heinz Dieter (1975). Zum Stand des automatischen Textworterkennung. *Grundlagenstudien aus Kybernetik und Geisteswissenschaft,* 16, 29–40.

Markarian, Ralf (1964). *The Educational Value of Esperanto Teaching in the Schools.* London: Universal Esperanto Association Research and Documentation Centre.

Noss, Richard (1967). *Higher Education and Development in Southeast Asia: Language Policy.* Paris: UNESCO.

Piron, Claude, and Tonkin, Humphrey (1979). *Translation in International Organization* Rotterdam: Universala Esperanto-Asocio.

Pool, Jonathan (1978). Elementary Models for Solving the Problem of Linguistic Diversity. Paper presented at the 5th International Congress of Applied Linguistics. Montréal, August.

Pool, Jonathan (1980a). The Economics of Artificial Languages: An Exploration in Cost Minimization. Paper presented at Fachtagung Angewandte Soziolinguistik, Universität Paderborn, June.

Pool, Jonathan (1980b). Lingva Egaleco, Lingva Malegaleco, kaj Lingva Diskriminacio: Problemoj de Konceptado kaj Mezurado. Paper presented at 65-a Universala Kongreso de Esperanto, Stockholm, August,

Pool, Jonathan (1980c). La Ekonomiaj Avantaĝoj de Planlingvo: Decidteoria Perspektivo. Paper presented at 3-a Internacia Konferenco pri Esperantologio, Stockholm, August.

Pratt, Terrence W. (1975). *Programming Languages: Design and Implementation.* Englewood Cliffs, N.J.: Prentice-Hall.

Rakuša, Rudolf (1970). *Metodiko de la Esperanto-Instruado,* 2nd edn. Ljubljana: Mladinska Knjiga.

Rubin, Joan, and Jernudd, Björn H., eds. (1971). *Can Language Be Planned?* Honolulu: University Press of Hawaii.

Rumbaugh, Duane M., ed. (1977). *Language Learning by a Chimpanzee.* New York: Academic Press.

Sherwood, Bruce Arne (1979). Schnelle textgesteuerte Sprachsynthese-Algorithmen für Esperanto, Spanisch, Italienisch, Russisch und Englisch. *Grundlagenstudien aus Kybernetik und Geisteswissenschaft,* 20, 109–121.

Stewart, William S. (1968). A Sociological Typology for Describing National Multilingialism. In *Readings in the Sociology of Language*, ed. Joshua Fishman (The Hague: Mouton), 531–545.

United Nations (1977). General Assembly. 32nd Session. *Report on the Implications of Additional Languages in the United Nations System.* Document A/32/237. 11 Oct.

United Nations (1979). *Statistical Yearbook*, 1978. New York: United Nations.

Wells, John (1978). *Lingvidtikaj aspektoj de Esperanto.* Rotterdam: Universala Esperanto-Asocio.

JONATHAN POOL

THE ECONOMICS OF ARTIFICIAL LANGUAGES

-abstract-

Proponents of artificial languages as media of communication among persons and organizations that do not share competence in any natural language often argue that a decision to use an artificial language for this purpose would confer economic benefits on all parties. In spite of this claim few individuals and organizations have ever decided to communicate across natural-language boundaries via an artificial language. This may mean that the claim is incorrect, that it is correct but not widely understood to be correct, that it is correct for coordinating sets of actors but not for actors in situations without coordination, or that decisions about language learning and language use are made on the basis of noneconomic as well as economic criteria. By applying polica-analytic and game-theoretic methods, we can determine the conditions under which it would be econimically advantageous to choose an artificial language as a medium of communivation.

SONDERDRUCK

Matthias Hartig (Hrsg.)

Angewandte Soziolinguistik

gnv Gunter Narr Verlag Tübingen

INHALTSVERZEICHNIS

(Fortsetzung Inhaltsverzeichnis siehe 3. Umschlagseite)

(Fortsetzung Inhaltsverzeichnis)

CIP-Kurztitelaufnahme der Deutschen Bibliothek

Angewandte Soziolinguistik / Matthias Hartig (Hrsg.) — Tübingen: Narr, 1981.
(Tübinger Beiträge zur Linguistik; 170)
ISBN 3 - 87808 - 170 - 7
NE: Hartig, Matthias [Hrsg.]; GT