3-30-1978

Translation Sciences Institute Slide Presentation Script

Roydon Olsen

TRANSLATION SCIENCES INSTITUTE SLIDE PRESENTATION SCRIPT

Roydon Olsen

(Note: This presentation is accompanied by a series of slides.)

The translation challenge the Church faces is tremendous. There are roughly 3,500 world languages. If we were to provide basic Gospel materials in all of the languages, and produce them at the present rate it would take 875 years to do so. The Church has created a list of approximately 250 priority languages which account for about 80% of the world's speakers. Producing translated materials for these languages, at the present rate, would still take over 60 years. We simply have to produce translations more reapidly and more efficiently.

Men of great vision have foreseen the role that modern technology would have in meeting that challenge.

Moreover, the beginnings will have been made in the automatic and instantaneous translation of languages, enabling people to understand one another across the barriers of Babel. (David Sarnoff)

Remember the message of the Master, "Go into all the world and preach the gospel to every creature"? Sarnoff says science will provide the vehicle and make it possible to do just that, because we can stand in Salt Lake and talk to all the world, and regardless of their languages they will understand what we say and thus will the world become prepared for the coming of the Son of God... My brethren and sisters, be prepared. All of this will happen. But before it does and while it is happening we are going to have some world-shaking events." (President Hugh B. Brown, "Pre-school Conference Address", BYU, Sept. 11, 1961.)

I believe that the Lord is anxious to put into our hands inventions of which we laymen have hardly had a glimpse.... He will open the gates and make possible the proselyting. Of that, I have great faith.... (President Spencer W. Kimball, Ensign, October, 1974)

The fulfillment of those prophecies is the mission of the Translation Sciences Institute in partnership with the Church Dept. of Translation and Distribution.

Since the founding of the Institute at Brigham Young University, we have received support, encouragement, and inspiration from the Church.

In 1972, then President, Harold B. Lee admonished TSI's director, Dr. Eldon Lytle, not to become discouraged, saying, "if you will make the Lord a partner, your computer translation project cannot fail."

In a presentation to the Board of Trustees, President Kimball, as well as President Ezra Taft Benson, reaffirmed the Church's faith in the project and emphasized their high expectations in the program.

President Kimball renewed the charge in his 2nd Century Address to Brigham Young University.

> BYU should become the acknowledged language capital of the world... We look forward to developments n your computer-assisted translation projects.... (President Spencer W. Kimball, BYU second century address)

Because of the strategic importance of the TSI project and on the specific recommendation of Elder Robert Hales, (who at that time was over translation and distribution), Dr. Ernest L. Wilkinson recently donated a one half million dollar IBM 370/138 computer to the Translation Sciences Institute. This gift is having a manifestly positive impact on the progress of computer translation.

At the time of this gift, a board was created with representation from the office of the Church Commissioner of Education, BYU, TSD, and Materials Management to coordinate BYU research as it applies to the Church.

By the end of 1978, the Department of Translation and Destribution's supplemental funding support for our research and development over the past few years will have surpassed $600,000.00. This funding has been vital in assuring the steady and continued development of TSI's work. Currently TSD funding represents 40% of our total budget. We are totally committed to showing the Church and BYU a return on this investment and look forward to continued success and cooperation.

PhaSE I PRESENTATION

Our translation research at the Institute is divided into two phases.

Phase I deals with computer text processing.
Phase II consists of our interactive computer translation program.

Phase I - Text processing

Text processing refers to the capability a computer has to automatically perform tasks formerly done manually by skilled typists, proofreaders, typesetters and printers.

Computer text processing technology is currently in wide use in journalism, industry, and within the Church. TSI has done extensive programming and development to make a text editor suitable to the work of translation, a field where it is still relatively unused.

Traditionally the translation process follows a large number of tedious steps, including rough drafts, several reviews and typings, and proofreading. Experience has shown us that before the manuscript is ready for publication it could be retyped as many as twelve times. Not only is this time- and resource-consuming, but new errors may be unintentionally introduced at any point, expecially if the typist is not totally familiar with the language is question.

TSI has taken the work station and developed special features especially needed by the translator. We have developed a custom-made translator editing system for the Church which enables the translator to draft, proofread, edit and correct his translation all in one step, without having to retype his copy. The computer can add, delete or change a letter, word, phrase, sentence or paragraph at the translator's command. It can format the text giving it proper spacing, even margins, page numbers, indentation, and other special features. TSI has done extensive software development in special formatting features, diacritical marks for foreign languages and typesetting routines for various languages. We use terminals connected to the Wilkinson computer or small, self contained units such as the Hewlett Packard work station.

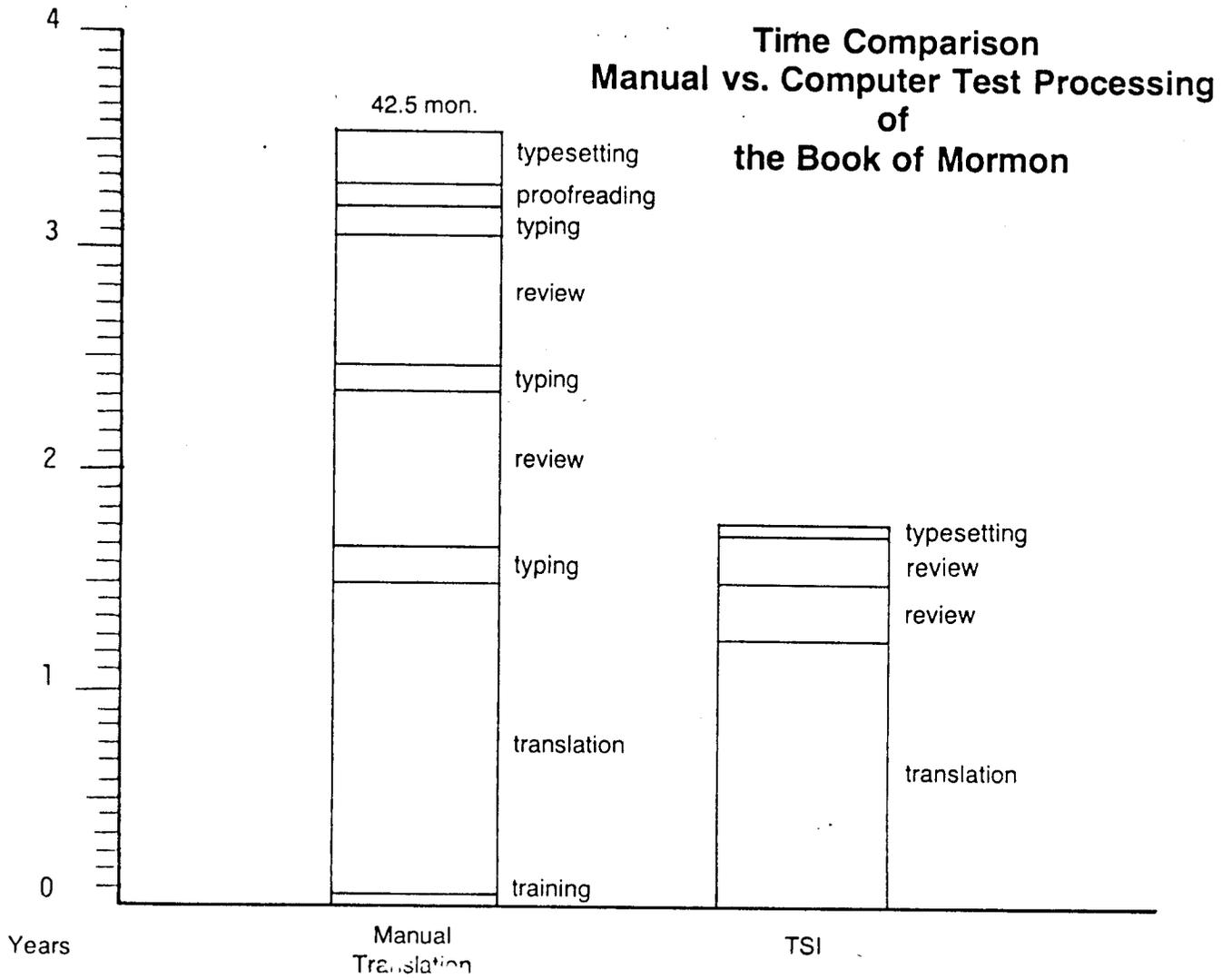Our Text Processing programs have enabled the Church to produce the following major works:

-Aymara Book of Mormon
-Topical Guide to the Scriptures

TSI text processing is currently in use in the following Lamanite languages:

1. Aymara   3. Quiche   5. Quechua
2. Cakchiquel   4. Quichua

We have been able to reduce significantly the time and cost of Book of Mormon translation in these languages:

Time Comparison
Manual vs. Computer Test Processing
of
the Book of Mormon

42.5 mon.

Manual Translation bar (bottom to top): training, translation, typing, review, typing, review, typing, proofreading, typesetting

TSI bar (bottom to top): translation, review, review, typesetting

Years

Manual
Translation

TSI

Cost Comparison
Manual vs. Computer Text Processing
of
the Book of Mormon

We have the capability now to aid Church translation in many of the current emerging languages with our automatic text processing programs on the Wilkinson computer. Without investments in any new equipment we could accomodate up to twelve languages by going to double shifts as outlined by the BYU and the Church at the time Dr. Wilkinson made this gift.

Of the approximately 250 priority languages, about half make use of a romanized alphabet for which we now have the capability to process up to the printing stage. Development of a special print wheel at a cost of under $3,000 would enable BYU Press to print any of the 125 languages.

With research in the area of special characters we could provide text processing for an even greater number of languages. Work is already underway in Chinese, and a major accomplishment has not been achieved with the capability to print 10,000 Chinese characters using a Versatec printer.

We now have the technology to adapt a work station for foreign language and install it in the field. Text processing systems which are compatible with Phase II computer translations will be needed at regional centers for post editing purposes. We look forward to working together to help put this valuable tool into the hands of our Church translators throughout the world.

PHASE II

TSI/BYU INTERACTIVE COMPUTER TRANSLATION
A program uniquely suited for Church needs.

The main features that set the TSI translation system apart from other computer-assisted programs are:

1. Human interaction. Our translation is interactive because the machine and human work together combining the advantages of both. In order to get high quality computer translation we have to acknowledge the limitations of the computer and compensate for these with human intervention. It is essential in understanding the meaning of a sentence that a human interact with the computer.

For example, in the sentence, "he flipped the coin on the table," there are at least four possible meanings:

(1) He flipped the coin onto the table.

(2) He flipped the coin that was on the table (as
    opposed to some other coin).

(3) The coin was on the table when he flipped it.

(4) He was on the table when he flipped the coin.

A computer would have a difficult time determining
which of these meanings was intended. A human would
normally know from the context.

2. Universal Semantic Code. Once the meaning is deter-
    mined, there must be a way to represent that meaning
    in the computer.   The human brain appears to be
    wired the same for everyone. Even though we may
    speak different languages, at a conceptual level we
    all utilize a universal code for handling informa-
    tion. All men share many concepts; each language
    may label the concepts differently.

3. Junction Grammar.   The mathematical way we have of
    representing the semantic code is called Junction
    Grammar and was developed by our director, Dr.
    Lytle. It enables us to represent meaning in pre-
    cise, mathematical terms in the computer.   Going
    back to the example of the coin on the table, there
    would be a different Junction Grammar representation
    for each of the possible meanings represented by one
    ambiguous sentence in English.

4. One to many. We can analyze the English once and
    then produce translations in any number of languages
    for which we have synthesis programs.

5. Vocabulary.   We project a programmed vocabulary in
    excess of 30,000 word senses by 1979.

ADVANTAGES OF THE ONE TO MANY SYSTEM.

    Paraphrase translation (what most humans do) requires a
separate analysis and interpretation for each language being
translated. There are disadvantages to this approach. For
example, the original text is subject to many interpreta-
tions by non-native speakers of English; because of this,
various translations of the same text may not be consistent
with each other and may not say what the original said.

    These problems are overcome in interactive computer
Translation because there is only one analysis of the origi-
nal, ant that by a native speaker of English. All of the
various language versions will be consistent with each
other, and if a mistake is detected in one version, it can

be readily corrected in all versions.   There will also be an
accurate transfer of information because the computer forces
the human to analyze the meaning more closely, as in the
"flipped coin" example. The machine forces us to account
for every structure, every grammatical relationship.

ADVANTAGES OF HUMAN INTERACTION

     Paraphrase translation which is practiced by the major-
ity of human translators, follows the Diversity Theorem,
which, simply stated, is that each person speaks differ-
ently. Within any given language there are dialects and
subdialects, and within dialects each individual has his own
manner of speaking, or idiolect. Each translation reflects
the translator's unique dialect and personal language
habits. Just as all of us possess a unique fingerprint, it
now appears that we each have a unique "language print".

     Because of this, translation reviewers tend to change
the translation to agree with their own dialect and idio-
lect--a major problem in Spanish translation for Latin Amer-
ica because what a Mexican translates may not satisfy a
Peruvian. Often changes are made in translations for no
other reason than " I wouldn't say it that way." This ten-
dency in review to change translation solely on a basis of
personal preference could, and often does, lead to a series
of revisions which, if not controlled, will result in a com-
binatorial explosion - the "revised edition" syndrome.

     Paraphrase translation is characterized by information
loss, as well. The translator often interprets text in his
own way. He also may not be aware of certain information in
the source. When he synthesizes a translation information
may be lost. Analysis of human translation both in and out
of the Church has shown this to be a widespread phenomenon.

Consider the following examples:

     ENGLISH: A deacon must be morally clean and pure in
     thought. He would never violate the wisdom or be
     involved in drug abuse. - Bishop Vaughn J. Feather-
     stone, from "The ᴋᴑle of the Deacon."

     SPANISH BACK TRANSLATION: ...would never break the word
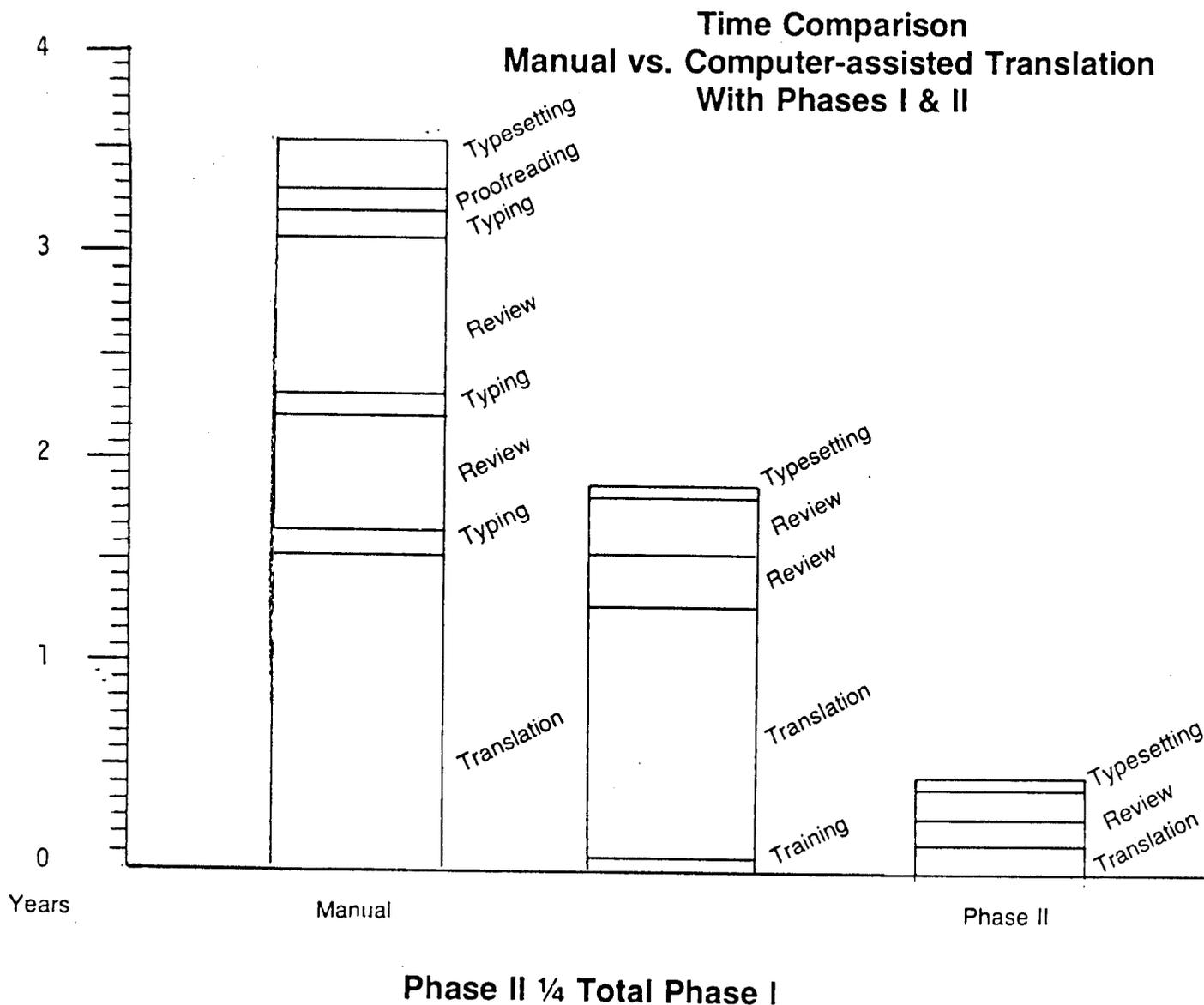   of wisdom or take drugs.

     How does Interactive Computer Translation help overcome
some of the traditional problems of paraphrase translation?
Interactive Computer Translation output does not correspond
to a given dialect or idiolect, but is intelligible to all
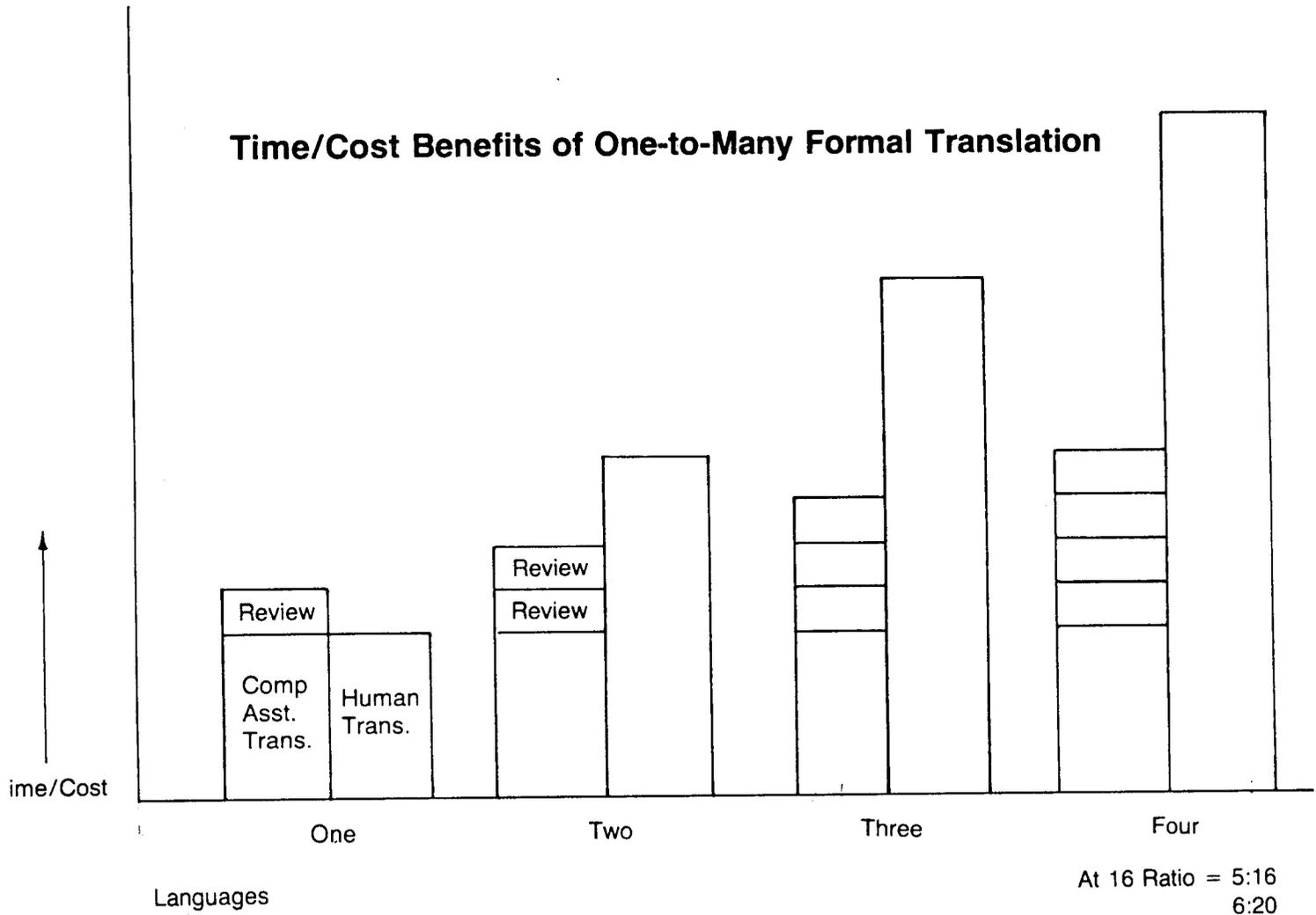speakers because of the input theorem, which states:

You understand more than you speak. That is, our
capacity to decode is greater .than our capacity to
encode. That is why even though we may speak a given
idiolect of a given dialect, our capacity to understand
bridges many idiolects, subdialects and dialects. So,
even though the computer translation output does not
correspond to a given dialect, it will be understanda-
ble. The Book of Mormon is a good example of this
principle. Even though it does not correspond to a
given dialect of English, it is understood by many.

Interactive Computer Translation will thus serve many
dialects and control the combinatorial explosion by supply-
ing one translation intelligible by all. The computer
translation will help control information loss because it
will be highly accurate. However, it will not be as
"pretty" as paraphrase translation. It will need some post
editing by a native of the target language to smooth out the
rough edges.

Time/Cost Benefits of TSI Computer Translation

Computer translation will be twice as efficient as
manual translation going into four languages. By adding the
advantages of both Phase I and Phase II, tremendous gains
can be realized. (See graphs on following two pages.)

**Time Comparison
Manual vs. Computer-assisted Translation
With Phases I & II**

Phase II ¼ Total Phase I

Time/Cost Benefits of One-to-Many Formal Translation

Current Status

We are on or ahead of schedule as we near 1979. Our
Phase II output was judged to be over 90% accurate by Sperry
Univac scientists. James Melnick, a consultant from Har-
vard, gave TSI the highest rating in his survey of machine
translation projects. We have made major breakthroughs and
our expectations are high that the project will attain the
desired level of performance.

Projections

The system will be ready for production by mid 1979 on
a limited basis. At that time the system will be "field
tested" and logistical and interface problems will be
resolved. Additional funding will be needed to continue
research to perfect the system and to add new languages

Conclusion: Interactive Computer Translation combines the
best of both worlds.

By taking advantage of those things which the computer
can do best and which the human can do best we can pro-
duce translation which is better, faster and more cost
efficient than what either could do alone.