Cognition, Information Processing, and Motivation

Edited by

Géry d'YDEWALLE

Department of Psychology
University of Leuven/Louvain
Belgium

1985

NORTH-HOLLAND
AMSTERDAM · NEW YORK · OXFORD

ISBN: 0 444 87783 5
ISBN set: 0 444 87790 8

Publishers:

ELSEVIER SCIENCE PUBLISHERS B.V.
P.O. Box 1991
1000 BZ Amsterdam
The Netherlands

Sole distributors for the U.S.A. and Canada:

ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
52 Vanderbilt Avenue
New York, N.Y. 10017
U.S.A.

Library of Congress Cataloging in Publication Data

International Congress of Psychology (23rd : 1984 : Acapulco, Mex.)
Cognition, information processing, and motivation.

(Proceedings of the XXIII International Congress of Psychology of the International Union of Psychological Science (IUPsyS), Acapulco, Mexico, September 2-7, 1984 ; v. 3)
"Published for the Internatonal Union of Psychological Science (IUPsyS)."
1. Cognition--Congresses. 2. Human information processing--Congresses. 3. Motivation (Psychology)--Congresses. 4. Memory--Congresses. 5. Learning, Psychology of--Congresses. I. Ydewalle, Géry d'. II. International Union of Psychological Science. III. Title. IV. Series: International Congress of Psychology (23rd : 1984 : Acapulco, Mexico). Proceedings of the XXIII International Congress of Psychology of the International Union of Psychological Science (IUPsyS), Acapulco, Mexico, September 2-7, 1984 ; v. 3.
BF121.I564 vol. 3 [BF311] 150 s [153] 85-10221
ISBN 0-444-87783-5 (U.S.)

PRINTED IN THE NETHERLANDS

Cognition, Information Processing, and Motivation
G. d'Ydewalle (editor)

THE MEAN LENGTH OF PROPOSITIONS IS 7±2 SYLLABLES - BUT THE POSITION OF LANGUAGES WITHIN THIS RANGE IS NOT ACCIDENTAL

Gertraud Fenk-Oczlon and August Fenk

University of Klagenfurt
Klagenfurt
Austria

As shown in a former study (Fenk-Oczlon, 1983), the mean number of syllables per simple declarative sentence translated into 27 different languages varies within a small range around 6.5 syllables. But the frequency distribution within this range can hardly be accidental: 15 languages fell into the class of 5 - 6 syllables; the frequency in the following classes was 7,3,2,0,1. The present study reanalyses data in order to test the assumption that the differing complexity of syllables is the decisive factor determining the position of a language within this distribution. The result of correlating the length of sentences (in syllables) with the length of syllables (in phonemes) was a coefficient of -0.77 ($p<0.1$). An additional finding: With SOV-languages translations proved to be longer than with non-SOV-languages. Presumably typologically different languages have developed in accordance with universal limitations of our processing capacity: Those with phonemically rich syllables carry more information per syllable and need equal time, but fewer syllables to express a certain proposition.

INTRODUCTION

As far as concerns linguistic universals and the role of memory in connection with language, there are substantially two lines of argument to be discerned: the first - decisively marked, no doubt, by Chomsky's views - considers language as a relatively autonomous organ common to all humans. Other cognitive functions, such as learning and memory, are according to this view, at best operative in the sense of disturbances of performance through a deficient functioning of the memory. The second line of argument, on the other hand, sees the cause of linguistic universals precisely in the general anthropological characteristics of the cognitive apparatus, its capacity limits and pre-linguistic interpretations of reality.

On the basis of the second line of argument, in an earlier experiment (Fenk-Oczlon, 1983), 22 simple declarative sentences (linguistic realisa-

tions of propositions) were constructed and were presented to 27 native speakers of different languages, who were asked to translate them into their own language and to determine the length of the translations given in terms of words and syllables. As expected, the length of sentences measured in syllables showed a small standard deviation, or in other words: the number of syllables used in very different languages in order to express propositions was relatively constant. It was on average 6.43 syllables per predicate-argument-structure (simple declarative sentence); the mean values per language varied between 5.5 and 10.2 syllables, i.e. roughly in the area around the oft-quoted "magical number 7". However, what was most noticeable in these results was the pronounced asymmetrical distribution of the languages around the mean value.

Table 1
The mean number of syllables/sentence in 28 different languages. (Fenk-Oczlon, 1983)

5-5.99	6-6.99	7-7.99	8-8.99	9-9.99	10-10.99
Dutch 5.05	Turk. 6.5	Ital. 7.5	Anjang 8.2		Japan. 10.2
French 5.3	Alban. 6.5	Greek 7.5	Korean 8.2		
Chin. 5.4	Port. 6.6	Span. 7.9			
Czech 5.4	Pers. 6.6				
Slov. 5.5	Hindi 6.7				
Hebr. 5.5	Panj. 6.7				
Germ. 5.5	Mazed. 6.95				
Icel. 5.5					
Eston. 5.7					
Russ. 5.7					
Sbcr. 5.8					
Engl. 5.8					
Ewondo 5.8					
Hung. 5.9					
Arabic 5.9					

The aim of the present study is, by means of a reanalysis of the data material, to learn something of the conditions which led to this asymmetrical distribution and are responsible for the position of an individual language

within this distribution. In thus, the hypothesis is made (which urges itself as a result of Table 1) that the syllable complexity is the decisive factor. Languages (e.g. Englis, Dutch, German, Chinese) with a high proportion of complex syllables (CCVC, CCCVC, CVCCC etc.) seem to prevail in the class 5.00 - 5.99 syllables, and languages with a high proportion of CV-syllables (e.g. Greek, Italian, Anjang, Japanese) at the other end of the range.

Other authors have already also expected differences concerning length of syllable and number of syllables per utterance, between languages which occupy extreme positions within our distribution. For example, Beckmann's (1982, p.134) explanation for the fact that Japanese sounds staccato in comparison to English was "that Japanese utterances will probably have more syllables than the English utterance, because syllable structure is simpler."

METHOD AND RESULTS

The assumption that the differing complexity of syllables is the decisive factor was tested by determining the number of phonemes per syllable in the translations given. In languages where our own knowledge on this issue was insufficient, native speakers were asked again where possible. In two languages (Hebrew and Arabic), this has not yet been possible.

The data so obtained, namely the average number of phonemes per syllable within an individual language, was correlated with the average number of syllables per sentence. The result was a correlation of $r=-0.77$ ($n=26$, $p<0.1$). That is to say for cross-linguistic comparison: the higher the mean number of syllables per simple declarative sentence, the lower the mean number of phonemes per syllable. This can be seen as a clear confirmation of our hypothesis.

An additional assumption was that typological differences in respect to agglutinating, fusional or isolating morphology might be relevant. The results of a comparison of SOV-languages (i.e. languages which prefer the word order Subject > Object > Verb) in our study - which show a tendency to open syllables and to be agglutinative in morphology (Lehmann, 1978) - with non-SOV-languages show that simple declarative sentences of SOV-languages ($n=7$) are with an average length of $\bar{x}=7.3$ syllables longer than those of SVO-languages ($n=19$, $\bar{x}=6.2$) or VSO-languages ($n=2$, $\bar{x}=5.7$).

DISCUSSION

The first requirement for high informational content is a large repertoir of possibilities (in our case, phonemes). The message or part of message - for instance syllables - composed of these elements are the more redundant (and therefore of lower informational content), the more the elements differ in their frequency and the more restricted the transitional possibilities are. These possibilities of composition are often referred to as the phonotactic possibilities of a language. If the phonemic inventory and the freedom of combination are given, then the length of the segment (e.g. syllable in phonemes) decides how much information this segment contains. Inversely, the larger number of elements per group is a precondition for a greater level of freedom in the combination possibilities. It can therefore be assumed that languages with a larger number of phonemes per syllable, generally carry more information per syllable and need fewer syllables to decode a message. Japanese occupies an extreme position in our distribution for good reason. It has a small phonemic inventory ($\sim$ 19 phonemes), almost exclusively CV-syllables, and thirdly decisive restrictions in the combination possibilities - each syllable therefore carries little information, more syllables must be used to decode a message.

If our information processing capacity is limited, a natural language which has developed in accordance with these limitations must grant to the perceiver (and the producer) more time to process syllables with higher informational content. It corresponds closely with a relatively constant information flow in language adjusted to our limits (Fenk & Fenk, 1980) that the informationally rich syllables in general require more phonemes and more time and that a unit of meaning (proposition) can be transmitted by a small number of informationally rich syllables, or a large number of informationally poor syllables.

The essential argument to be drawn from our results is that the "magical number seven" (Miller, 1956) is a surface-phenomenon only. Capacity limits of our cognitive processes are at best and in the most general form to be defined in terms of subjective information. These limits are localized in a region, which - in some cases - is equivalent to the amount of subjective information that is carried by approximately seven elements or seven arrangements (of elements) with a very high internal redundancy, i.e. with high subjective conditional probabilities between the elements constituting an arrangement. (The accurate number of elements or chunks of the hierarchi-

cal level n, composing a chunk on level n + 1, depends on the subjective informational content of level-n-components). And in languages the arrangement with the highest internal redundancy precisely is the syllable, a fact that should be added to the list of facts that according to Mehler, Dommergues, Frauenfelder, and Segui (1981, p.298), "led some.....to endorse the syllable as the basic segmentation unit."

REFERENCES

Beckmann, M. (1982). Segment duration and the 'mora' in Japanese. Phonetica, 39, 113-135.

Fenk, A., & Fenk, G. (1980). Konstanz im Kurzzeitgedächtnis - Konstanz im sprachlichen Informationsfluss? Zeitschrift für experimentelle und angewandte Psychologie, 27, 400-414.

Fenk-Oczlon, G. (1983). Bedeutungseinheiten und sprachliche Segmentierung. Tübingen: Narr Verlag.

Lehmann, W. P. (1978). English: A characteristic SVO language. In W. P. Lehmann (Ed.), Syntactic typology. Sussex: Harvester Press.

Mehler, J., Dommergues, J. Y., Frauenfelder, G., & Segui, J. (1981). The syllable's role in speech segmentation. Journal of Verbal Learning and Verbal Behavior, 20, 298-305.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.