# JUBA - A SHORT INTRODUCTION

Between the years 1980-1983 the Department of Slavonic Languages at the University of Lund has carried out a linguistic research project entitled "The Development of the Family Language of Serbocroatian/Croatian Speaking Children aged 4 - 18 in Sweden", abbreviated JUBA (from "jugoslaviska barn" Yugoslav children). The project was financed by two Swedish research foundations (the governmental HSFR: the Swedish Council for Research in the Humanities and Social Sciences and RJ: the Bank of Sweden Tercentenary foundation).

1.1. During the sixties and early seventies, about 60.000 workers from Yugoslavia immigrated into Sweden, 81% of them with Serbocroatian/Croatian as their native language (in the following texts, the abbreviation S/C will be used). They obtained the same legal status as Swedish citizens in terms of the Swedish civil law and were offered the option of retaining their Yugoslavian citizenship or becoming Swedish citizens. Their children have been guaranteed a certain amount of school instruction in S/C by teachers whose mother tongue is S/C. For training of those teachers a special program has been started at the Schools of Education in Malmö, Göteborg-Mölndal and lately in Stockholm.

1.1.2 The main flow of this immigration ceased in 1972 and consequently about 45000 citizens of Yugoslav origin are now living in Sweden, dispersed throughout its various industrial centres. Since the main group of immigrants were at the child bearing age, a generation of children is now being born in Sweden to families where one or both parents speak one of the "Yugoslav" languages as their first language. Their children, born in Sweden thus acquire their parents' language (referred to as "family language", "hemspråk") in circumstances radically different from those of the children living in Yugoslavia. This situation is described in the paper Stankovski - Ďurovič - Tomašević, Development Structures in the Family Language of Jugoslav Immigrant Children in a Swedish Language Environment (pp. 11ff.).

1.1.3. JUBA's aim has been to investigate what happened to the language of the S/C speaking Yugoslav children, both of those who arrived in Sweden with their mother tongue acquired (presently aged 15 or older) as well as of those who have acquired their mother tongue in circumstances of diaspora (see Stankovski, 1981, Prvi jezik druge generacije). We see a possibility of tracing the

language processes either in their tendency to constitute a more or less correct S/C, or in the various forms of its collapse, by following the language development of a selected group of second generation children.

JUBA a) gathers language records of S/C speaking children aged 4 - 18
b) gathers a certain amount of social background data and sociolinguistic information about the informants
c) makes efforts to follow the same speaker over several years (several year-cuts).

2.1. The object of investigation is the formants' spoken (not written language.) Each test consists of the following parts:

a) a S/C story retold both in S/C and in Swedish (S/C stimulus)
b) a Swedish story retold in S/C and in Swedish (Swedish stimulus)
c) a picture to be described in detail in S/C (non-verbal stimulus)
d) a story about a subject chosen by the informant and told in S/C (narrative without stimulus)
e) a lexical test, consisting of 50 simple coloured pictures to be named in S/C and Swedish
f) directed morphological tests (case forms and numerals)
g) phonological tests (realization of phonemes, identification of speech sounds)
h) every informant is asked 103 questions about his or her personal background and about his or her language behaviour.

In accordance with Swedish rules for research ethics in the humanities the informants are identified by a number code only and remain, consequently, anonymous as persons.

This approach allows for one informant's record to yield about 500-1200 S/C words (100-250 predications) in continuous narrative texts, 38 grammatical forms, a maximum of 50 S/C and 50 Swedish words naming 50 given referents, about 500 Swedish words in the same texts, some directed information about pronounciation of selected S/C sounds (mainly affricates and sibilants). This information is recorded on magnetic tapes.
Up to this point, tests from 547 children have been gathered, 168 of them have been tested twice, 8 three or more times.

The S/C texts are subsequently transcribed, adapted and stored in computer memory.

2.2.1. As a rule, the transcription is orthographic: in order to enable computer processing, the S/C Latin orthography is used. Corrupt or incorrect sounds are written as they sound (e g lisisa). Conventional punctuation is not used; instead the text is segmented into predications (by means of +) and the sentence intonation contour is noted // (cadence, falling tone) and / (anticadence, rising tone). A segment delimited by // is called "a sentence". All pauses are noted by ..., all un-finished words are transcribed and noted by a dash - (e g ma-, magar-). All interventions by the tester are transcribed (in parentheses).

2.2.2. For internal purposes the corrupt sounds are reconstructed after an asterisk (e g lisisa*lisica). All "incorrect" grammatical forms (from the standpoint of the S/C standard) are reconstructed after the sign & (e g od lisicu&lisice). The signs * and & can be combined, e g od lisisu*lisicu&lisice. The sign * is also used to reduce all reflexes of jat to one common form ě (e g čovik*člověk, čovek*člověk or čovjek*člověk). In the texts published here (pp. 151ff) these reconstructions are omitted: only the physically present text is reproduced.

2.2.3. Other nonsegmental properties of the physical texts, such as vowel length, the S/C tonal accent, prominence, different kinds or rhematization etc are not transcribed for the time being. Neither have the texts in Swedish been transcribed. In this way, a large volume of valuable data can be extracted from the tapes and transferred to transcripts for further investigation.

2.3. The tapes, the transcripts as well as tests with the testers' notations attached, are in archives at the Slavonic Languages Department in Lund.

3.0. All the transcribed texts are stored in the computer memory in the Lund University Computer Centre (LDC).

3.1.1. For handling of the stored material a system of programs has been constructed by Ingemar Dahlstrand, LDC (see his article, pp. 107ff).

3.1.2. The computer storing has enormous advantages. It guarantees errorless and always identical written output at all times and further yields the possibility of computer setting. Its most important advantage is, however, the possibility of automatic

operations on the stored texts. At this point JUBA has at its disposal programs for searching in the corpus and for automatic grammatical analysis. By means of SPSS statistical operations can be carried out on any part of the stored data. A more detailed account of the programs is given in the article by I. Dahlstrand.

4.0. Some results of the linguistic investigations in JUBA are published in this volume of Slavica Lundensia.