BABEŞ-BOLYAI UNIVERSITY

CLUJ-NAPOCA

FACULTY OF LETTERS

ACQUISITION OF ROMANIAN LANGUAGE AS A SECOND LANGUAGE.

A1-LEVEL INTERLANGUAGE

Doctoral Thesis

ABSTRACT

DOCTORAL SUPERVISOR:
PROF. EMERITUS G. G. NEAMŢU, PH.D.

PH.D. CANDIDATE: LAVINIA-IUNIA VASIU

2017

TABLE OF CONTENTS

Foreword

Key-words: interlanguage (IL), natural/first language (L1), foreign/second language (L2), target language, second language acquisition (SLA), first language acquisition (FLA), competence, performance, learner corpus (LC), level of proficiency, A1 level, transfer, input, output, variability, pragmatic interlanguage, error analysis (EA), obligatory occasions analysis (OA), frequency analysis (FA), functional analysis (FcA), communicative strategies, strategic competence, linguistic knowledge, negative distinguishing linguistic features (NSF), positive distinguishing linguistic features (PSF), L1 specific features (L1F), transitional features (TF).

In the case of many languages, be they perceived as 'small' languages or international languages, the number of studies based on learner corpora (LC) is continuously growing (starting with the 1980's), whereas the relevance of the data offered by them in domains such as applied linguistics, didactics and second language acquisition (SLA) is unanimously acknowledged. Nevertheless, in what regards Romanian as a foreign or second language (L2), such studies relate to rather isolated areas, being generally grounded on rather too small corpora in order to be considered significant. Taking all these into account, our research proposal seems to be more than welcome.

To start with, the originality of our approach lies in the proposal of a corpus on every level (although not extensive – approximately 70,000 words), devised on the oral productions of the non-native speakers of Romanian language (CORLS), originating from various cultural and linguistic environments. Moreover, there has been no similar attempt so far. At the same time, such a CORLS analysis on more levels (morpho-syntactic, functional/pragmatic, lexical, of the communicative strategies) and by different methods (error analysis, obligatory occasions analysis, frequency analysis, function-form analysis, etc.), as well as the categorisation of the data obtained as *specific linguistic features* (SF) for the interlanguage specific to A1 level (IL = a structured linguistic system, situated on the verge between L1 and L2), are highly relevant in further domains.

Our research is based on various treaties, studies and articles claimed by applied linguistics, first and second language acquisition (SLA), didactics and learner corpus (LC), elaborated during 1960-2017, mostly in the Anglophone areas. Major works belonging to L. Selinker (interlanguage), S. P. Corder (error analysis), R. Ellis (SLA) or several articles and studies elaborated by certain celebrated researchers such as V. Cook, H. Dulay, M. Burt, S. Krashen, S. Gass, D. Larsen-Freeman, G. Leech, B. North, M. Pinemann, E. Tarone, etc., represent the theoretical basis of the present thesis, whereas the idea of the entire project owes it to the volume *Criterial Features in L2 English. Specifying the Reference Levels of the Common European Framework* (J. A. Hawkins, L. Filipović). The relatively recent and rather timid preoccupation for such a domain, from the Romanian academic area, explains the penury of a significant number of bibliographical references. We feel bound therefore to mention E. Platon, L. Avram and, as part of the new generation, C. Bocoș, Ș. Tărău, A. Arieșan, authoring certain relevant articles and studies. At the same time, a present-day crucial document for Romanian as L2, *The Minimal Description ...*, has been correlated (together with *The Common European Framework of Reference for Languages - CEFR*) with the results of the research, in the second part of our thesis.

In what regards the content organization, the present study is structured into two sections: the first section, visibly more dense theoretically and bibliographically, delineates the theoretical frame of the research, while the second section, having a rather practical and novel character, presents the results of the analysis. The two sections are followed by the general conclusions and the annex represented by CORLS (which also includes certain meta-data in the first pages).

Starting from the SLA theories and comprising four chapters, **Section I** outlines the entire research. **The first chapter** brings arguments for the necessity of a research similar to the one we have proposed in the Romanian area. The most relevant documents for the modern teaching/learning and assessment of Romanian as L2 are here under examination. The *Common European Framework of Reference for Languages* (*CEFR*), the work whose issuance triggered the European linguistic phenomenon of the six main levels in the acquisition of any foreign language (A1, A2, B1, B2, C1, C2), proposes functional descriptions, in highly general descriptors, not based on learner-corpus studies – currently considered, in fact, the entirely valid sources for such descriptions. *The Threshold Level* represents the first attempt to provide detailed descriptions of the *CEFR* applied to Romanian (B1 level)

and one of the documents lying on the basis of elaborating *The Minimal Description of Romanian Language* (*A1, A2, B1, B2*), a work proposing inventories of lexical elements, communicative functions, grammatical categories, etc., for the first four levels broadly defined in the *CEFR*. This description differs from certain ones elaborated on other languages due to the fact that its basis is not represented by the theories on second language acquisition (SLA) or the theories resulted from corpus linguistics (CL), but by the authors' intuition and experience, as well as other existing similar documents. Our study starts exactly from the previously mentioned drawback and initiates the realization of a description of the interlanguage (= the language of the non-native speakers), A1 level, founded on the results of a corpus analysis on the same model in the well-known project *English Profile*. The "profile" of Romanian as L2 (IL on A1-level) will be thus outlined by extracting certain distinguishing linguistic features concerning the communicative strategies, the linguistic, sociolinguistic and pragmatic knowledge, categorized by the analysis of a corpus devised (especially for such a goal) out of the oral productions of the non-native speakers of Romanian, situated on the first level of linguistic competence.

Focusing on the terminological and conceptual definitions, **the second chapter** explains SLA as the acquisition of any language but the natural/first language (L1), while the term *acquisition* is alternately used with *learning* throughout our thesis. In a *CEFR* perspective and Bachman's point of view, the *competence* includes both linguistic knowledge and the strategic competence, whereas the SLA theories aiming to define it are classified according to the direction and approach, being roughly delineated in this chapter, which highlights the essential aspects of every theory: the input, the explicit/implicit knowledge, the output, variability, the natural order of acquisition, the transfer, the training in L2, etc. We also mention here the general directions of the behaviorist and structuralist theories, of the *Monitoring Theory*, of the generativist theories (*Full Access/No Transfer*, *Full Access/Full Transfer*), of the cognitivist approaches (the *Competitional Model*, the *Operational Principles*, the *Processibility Theory*, of the *Complexity Theory*, the *Interactional Theory* and so on), and, finally, the ideas triggered by such theories are systematized and chronologically ordered with a view to capturing a broad picture of the study evolution in SLA.

Chapter III of the first section details the issue of IL, its analysis methods, as well as the processes and factors implied in its development. We operate with L. Selinker's definition, by which IL is a structured linguistic system, functioning on its own principles that do not entirely belong to the first language, neither to L2, being in a continuous process of restructuring, so as to draw close to the L2 system. The special features of this peculiar linguistic system are: *permeability* (the new forms, derived from the input, output or the generalization/simplification of certain forms in L2, easily permeate it), *variability* (synchronic and diachronic, systematic – determined by the communicative context – and non-systematic/free) and the *systematic character* (irrespective of the variability, in any moment of its evolution, IL can be described as a logical and coherent linguistic system). In what regards the analysis methodology of IL, our focus is the error analysis (EA), following Corder's methodology: error identification, error description, explanation and their ultimate assessment, a methodology we also employed (excepting the last stage) in the analysis of the morpho-syntax. Nevertheless, the drawbacks of the EA have been completed by certain aspects of the obligatory occasions analysis (OA) and the frequency analysis (FA). In order to extract the IL pragmatic features, the functional analysis (FcA) has been applied, of the *function-form* type. Among the essential factors taken into consideration in our analysis, there are the *transfer*, the *input* (having a crucial importance in the acquisition, yet reduced, in our study, to the interlocutors' interventions – examiner/other candidate – and to the requirements of the tasks to be solved) and the *natural order of acquisition* (in what regards this aspect, our conclusions are rather timid, since it is not a longitudinal study to allow the formulation of truly valid hypotheses). The last part of the chapter illustrates the manner in which these concepts are made operational in our research.

Chapter IV, the last chapter in Section I, defines the concept of *corpus* and briefly presents the existing corpora on the Romanian language. Nevertheless, in contrast with other languages, no corpora on the Romanian language as L2 have been elaborated so far. Such learner corpora (LC), containing productions of non-native speakers, are extremely relevant both from a linguistic perspective (corpus linguistics – CL, applied linguistics – first and second language acquisition), and a didactic point of view (elaboration of curricula, syllabuses, manuals, dictionaries and other documents, assessment tools and various materials used for the training of didactic staff, etc.). Consequently, we consider that our study, based on a learner corpus for the Romanian language, comes at the right moment. CORLS

contains oral productions of various types (a dialogue with a native speaker, a monologue, a transactional dialogue, with a non-native speaker), on different topics, collected from 172 students, coming from several linguistic environments, during the oral test of the Examination for Romanian as a foreign language, A1 level, from three different exam-sessions. The transcription of the productions also includes the marking of certain phenomena such as the wrong pronunciation of the words, the empty and full pauses, interruptions and self-interruptions, the use of syllables and the emphasis, the final and non-terminal ascending and descending intonation and so forth – thus, all phenomena specific to the spoken language and, to a greater extent, to a spoken L2, A1 level. We consider that, by the fact that the transcription does not neglect such aspects, CORLS represents a useful tool, being able to be analyzed from more perspectives, with various goals. In order to segment the productions, we propose a syntactic unit of speech analysis - ASU –, the way it is defined by Foster and his collaborators: „an independent sentence (be it also elliptical) and its dependent sentences", offering further specifications on segmentation in problematic situations (interruptions, self-interruptions, one-element utterances, etc.). This chapter also tackles on the "attackable" points of the proposed corpus, namely the dimension and the degree of representativeness.

Section II, as previously mentioned, highlights the results of the CORLS analysis, on each of the three main coordinates describing the competence: the *communicative strategies* (the strategic competence), the *sociolinguistic and pragmatic knowledge* (the communicative functions – pragmatic IL), *linguistic knowledge* (the morpho-syntax and the lexis).

Thus, in **Chapter I** of the second section, we introduce the distinguishing linguistic features regarding the communication strategies used by A1 learners of Romanian as L2. Even though the CEFR does not provide any descriptions concerning this aspect for level A1, the CORLS analysis demonstrated that learners use such strategies even at the very first level of language proficiency. Consequently, it appears that whenever they have difficulties in articulating the message, learners frequently draw on code-switching, indirect request for help or approximation, when they need extra time for planning, they fall back on repetition or non-lexical fillers, whereas learners who speak Romance languages frequently use *I think that.../I don't know...* with the same purpose.

The second chapter focuses on the communicative functions encountered in CORLS and on their linguistic expression. A functional analysis of the function-form type revealed the fact that learners tend to use some constructions as fixed, *unbreakable* ones when trying to express different functions, on the one hand, and that many errors occur when they attempt to replace one element of those formulaic chunks in order to integrate them in new contexts, on the other hand. The analysis shows that at this level learners introduce themselves or someone else using the verb *to be*, they thank people using *Mulţumesc!/Mersi! (Thank you!/Thanks!)*, and they reply with *Cu plăcere! (You're welcome!)*. Learners justify using *pentru că* (*because*) and they can express uncertainty (*cred că/probabil – I think that/probably*), affinity and animosity (the verb *a-i plăcea – to like*), preference *(A prefera + verb/substantiv – to prefer + verb/noun)* and so on.

In **the third chapter** we investigate the lexis on two coordinates: volume and control. Thus, we propose frequency lists for the adjective, the verb and the adverb, but we also indicate the drawbacks of the lexical elements used in CORLS, obvious in the frequent omissions, confusions, lexical inventions, code-switching, etc.

The chapters IV-VI reveal the distinguishing features regarding morpho-syntax. Thus, with respect to the noun, the CORLS analysis demonstrated that, leaving aside some exceptional cases that can be explained by the input frequency – the definite article attached to the noun in the possessive construction or in the position of instrumental or associative complement, for example –, the categories of number and determination do not constitute features that are correctly acquired at level A1. In a similar manner, the adjectival agreement and the verb conjugation in indicative and subjunctive (present) both seem to raise a lot of problems for L2 learners. Among the errors determined by negative transfer we have to mention the omission of *să* in the contexts that require the use of the subjunctive, in the case of the learners who speak Romance languages, and the frequent elision of the verb *a fi/to be* in Arabic speakers' utterances. The control of the reflexive pronouns and of the personal ones in dative does not constitute a distinguishing feature for the IL at this level. However, the personal pronouns in nominative, used as cohesive devices, as well as the interrogative, invariable *ce* are used correctly. Adverbs occur seldom, but when they do occur, they seem to be adequately integrated in the utterance.

The conjunctions *și/and*, *pentru că/because*, and *sau/or* are usually used correctly, whereas prepositions seem to be inadequately used in most cases.

Finally, the last chapter, which addresses the ***Conclusions***, reiterates in less detail the results presented in the previous chapters. We draw attention on the fact that our research confirms the results generated by similar analyses in the case of other languages as L2: transfer seen as a learning strategy at level A1 plays an insignificant role regarding the acquisition of those features connected to morpho-syntax; transfer seen as a communicative strategy is mostly reflected in code-switching and in literal translations; the input and the output have a huge impact in the order of acquisition, etc. Also, this chapter proposes a comprehensive scheme which attempts to shape the full picture of the IL at the A1 level on all three important dimensions defining the general linguistic competence: *the strategic competence, the sociolinguistic and pragmatic knowledge and the linguistic knowledge*.

By building such a detailed profile of A1 level interlanguage, we consider our study to be relevant firstly in the field of didactics (for developing curricula, L2 grammars and dictionaries, course books, tests, grids and scales, materials for teacher trainings, etc.) and secondly in that of linguistics, constituting a good starting point for future research.