

FOREWORD

The studies in this book originated from an international program called *Lexical Syntax in Basque* (1999-2001), which was funded by the Department of Humanities and Social Sciences of the CNRS (UMR 5478) and the Basque Autonomous Community's Ministry of Education and Research. As its title indicates, this program was oriented toward the study of relations between lexicon and syntax in Basque. The program assembled approximately about fifteen Basque researchers, both computer scientists and linguists, most of whom were associated with the University of the Basque Country. Half the papers consists of studies conducted in computational linguistics, and half result from more classical morphosyntactic studies.

The five articles written by computational researchers reflect a collective undertaking which has been developed in recent years by members of the IXA group in San Sebastian. One of this research group's primary goals is to build automatic analyzers of texts. The papers presented in this book, which are mainly devoted to verbal subcategorization, provide an insight into the manner in which this significant aspect of lexical syntax is dealt with in the field of corpus linguistics.

Two syntactic analyzers have been created for the parsing of Basque corpora. One (called PATRixa) uses a unification-based formalism, and it is dealt here by Aldezabal, Aranzabe, Atutxa, Gojenola and Sarasola, in the context of the automatic analysis of verbs' subcategorization; the other one is based on a Constraint Grammar (Karlsson 1995). The foundation of the latter is presented in this book by Aduriz & Diaz de Ilarraza in a paper devoted to the way morphosyntactic ambiguities in texts may be analyzed.

Furthermore, two other applications are presented which pertain to the extraction of information about verbal subcategorization. Arriola, Artola and Soroa show how dictionaries can be used as a starting point to guide lexical acquisition from corpora, while Aldezabal, Aranzabe, Atutxa, Gojenola and Sarasola examine how lexical information based on the argument *vs* adjunct distinction can be extracted from raw corpora, after the application of a partial parser and statistical filters. In the fifth paper, Aldezabal and Goenaga discuss several general problems regarding the way verbal subcategorization can be dealt with in computational linguistics, and they present the results obtained in their analysis of subcategorization patterns for one hundred verbs.

As indicated before, the other papers have a more classical flavour. The paper by Alberdi consists of a careful examination of the use of ergative case in Basque borrowed verbs, especially with non transitive verbs. He shows that Spanish unergative verbs

which are formally reflexives (Mendikoetxea 1999) are treated as absolutive intransitive verbs in Basque borrowings, and he demonstrates that in fact, the tendency to use an ergative case marking with intransitive borrowed verbs (Sarasola 1979) is restricted to non-reflexive (most of the time agentive) Romance verbs. The remainder of the inquiries deal with verb derivation. Artiagoitia's paper analyzes the case of *-garri*, a deverbal suffix which has been traditionally considered as having two opposite values (an active one and a passive one). Artiagoitia offers a unified account of the suffix, and dismisses the notion that *-garri* masks two homophones. Because the so-called active value of *-garri* appears with object-experiencer psych-verbs, he proposes an analysis which frames these verbs as lacking an external argument and having a derived subject (Belletti & Rizzi 1989). Such an analysis is not compatible with the one offered by Oyharçabal in his paper devoted to verbs having the causative alternation. In Oyharçabal's view, alternating causative verbs (including psych-causatives of the *frighten / preoccupare* type) project a *Cause* head in syntax (Pylkkänen 2002). Based on lexical decomposition, his papers examines the different types of lexical causatives in Basque, and concludes they can be characterized as verbs that express a change in the physical constitution, the location or the psychological state of the theme, patient or experiencer. Zabala offers a broader analysis of causation in her work. She analyzed this notion along with other semantic concepts such as agentivity, control and intentionality. Starting from data obtained in technical texts, she suggests checking predictions that concern semantics and morphosyntax of certain kinds of verbs, including deverbal derived nouns and adjectives. The study of Odriozola is devoted to a systematic classification of derived verbs: denominal, deadjectival and depostpositional (or deadverbial) verbs are analyzed, mainly using semantic criteria in each case.

In 2001, as the *Lexical Syntax in Basque* program was finishing, Professor Miren Azkarate, who was the head of the group of researchers from the University of Basque Country in this project, was called upon to assume new responsibilities and was no longer in position to contribute to this volume. Although she was unable to participate in this edition, my colleagues and myself would like to thank her for her commitment to in the completion of this project. We would also like to thank Joseba Lakarra for the help he has provided to us in the preparation of this volume.

In closing, it was during the correction of the proofs of the papers that we learned that Rudolf de Rijk had passed away. We knew him to be gravely ill, and like other Basque linguists, we were greatly saddened at this news. Basque syntactic studies owe so much to his work. We dedicate this book to his memory.

B. Oyharçabal