

**OSDL**

Ohio State  
Dissertations in Linguistics

# **Nominalizations in Basque: A Case in Language Attrition**

**Kutz M.C. Arrieta**

**1998**

For ordering information contact **OSDL** (Ohio State Dissertations in Linguistics):

OSDL  
Department of Linguistics  
222 Oxley Hall  
1712 Neil Avenue  
Ohio State University  
Columbus, OH 43210-1298  
U.S.A.

[osdl@ling.ohio-state.edu](mailto:osdl@ling.ohio-state.edu)

For dissertation abstracts and a listing of available titles visit our WWW server:  
<http://www.ling.ohio-state.edu/dissertations/>

NOMINALIZATIONS IN BASQUE: A CASE IN LANGUAGE ATTRITION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for  
the Degree Doctor of Philosophy in the  
Graduate School of The Ohio State University

Copyright by

Kutz M. C. Arrieta

1998

By

Kutz M. C. Arrieta, M. A.

\*\*\*\*\*

The Ohio State University  
1998

Dissertation Committee:

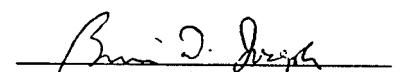
Professor Brian D. Joseph, Adviser

Professor Catherine Callaghan

Professor Amy Zaharlick

Professor Terrell A. Morgan

Approved by

  
Brian D. Joseph  
Adviser  
Interdisciplinary Graduate  
Program

## ACKNOWLEDGMENTS

I wish to thank my adviser, Brian Joseph, for his support and encouragement. I wish to thank also all the members of my committee: Cathy Callaghan, Terrell Morgan, and Amy Zaharlick. Without their support this would not have been possible.

I wish to thank my friends, especially Pat Pepe, who helped me throughout the dissertation writing process with her skills in documentation and her support as a friend. I also want to thank Eve Baron, who tried to make my written English acceptable. In general, I want to thank all my friends, including my husband, for their support.

## VITA

November 2, 1959.....	Born – Donostia, Basque Country, Spain
1982.....	M.A. Linguistics, Université de Bordeaux III
1985.....	D.E.A. Basque Studies, Université de Bordeaux III
1987.....	M.A. Romance Languages, The Ohio State University
1987 – 1993 .....	Graduate Teaching and Research Assistant, The Ohio State University
1993 – present.....	Computational Linguist, Logos Technology Center

## PUBLICATIONS

### Research Publication

1. "El lingüista y la traducción automática. Logos lingüísticos en un sistema de traducción automática." In: *Apuntes de SpanSig*, September 1994. New York.
2. "That's Something We Can Trace Back to the Complementizer" (with Richard Danford). In: *Issues and Theory in Romance Linguistics*, Michael L. Mazzola (ed), Georgetown University Press (1994).
3. "Sobre la pronunciabilidad de grupos consonánticos marginales: evidencia empírica del español" (with T.A. Morgan). In: *Hispanic Journal*, 11:2 (Fall 1990), 171-184.
4. "Nuevas aportaciones al estudio de los grupos consonánticos marginales en español" (with T.A. Morgan). In: *Anuario de letras XXVIII* (1990), 273-283 (Mexico City: UNAM).

LIST OF TABLES

<u>Table</u>	<u>Page</u>
Table 1. Comparison of Distribution and Properties among English, Catalan, Italian and Spanish.....	67
Table 2. Basque Population of the United States - 1990 and (1980).....	245

LIST OF ABBREVIATIONS

A	absolutive
Acc	accusative
adl	adlative
adv	adverbializer
asp	aspect
Aux	auxiliary
bare	bare form
cond	conditional
comp	complementizer/subordinator
D	dative
E	ergative
fut	future
G	genitive
ger	gerund
ind	indicative
inf	infinitive
instr	instrumental
imp	imperfect
loc	locative
neg	negative
nom	nominalizer
part	partitive
past	past
pl	plural
pple	past participle
pres	present
pret	preterite
prf	perfective
prog	progressive
prol	prolative
rel	relative
sg	singular
subj	subjunctive

4.5 Conclusion .....	235
<b>5. RESULTS AND DISCUSSION .....</b>	<b>237</b>
5.1 Introduction .....	237
5.1.1 Historical Background.....	239
5.1.2 Basques in the New York Area.....	244
5.2 The Speakers.....	247
5.2.1 Background .....	249
5.2.2 Age.....	250
5.2.3 Education .....	250
5.2.4 Speaking Abilities.....	251
5.3 The Data .....	252
5.4 Analysis .....	254
5.4.1 Basic features .....	255
5.4.2 Lexicon .....	258
5.4.3 Distribution of -te Nominalizations and Derived Nominals .....	260
5.4.4 Sentential versus Nominal .....	277
5.4.5 Control and Case Marking.....	278
5.4.6 Complexity.....	280
5.4.7 Code Switching .....	280
5.4.8 Attitudes .....	284
5.5 Conclusion .....	291
<b>6. CONCLUSION .....</b>	<b>294</b>
<b>ENDNOTES .....</b>	<b>298</b>
<b>APPENDIX A SPANISH SENTENCES GIVEN TO SUBJECTS &amp; BASELINES .....</b>	<b>300</b>
<b>APPENDIX B ENGLISH SENTENCES GIVEN TO SUBJECTS.....</b>	<b>312</b>
<b>BIBLIOGRAPHY .....</b>	<b>315</b>

## CHAPTER 1

### INTRODUCTION

In this study I investigate nominalizations in Basque in a language contact situation. I adopt as a working hypothesis that Basque nominalizations are a complex structure and that they are thus vulnerable in language contact situations in which simplified structures might be expected to prevail. The relevant notion of complexity is discussed in chapter 4. This hypothesis is tested against the data in Chapter 5.

My starting hypothesis also builds on the idea that language attrition, viewed as a variant of language change, can be diagnosed through studies of speakers' linguistic behaviors.

My approach to this problem is an Anthropological Linguistics type of approach. I did not focus on sociolinguistic data, rather on linguistic and anthropological data. I did not apply statistical methods, rather I looked at the linguistic content of the data.

concerned, however, with language attrition in a diaspora situation, the Basque community of the New York area.

## CHAPTER 2

### NOMINALIZATIONS, DERIVED NOMINALS AND NOMINALIZED INFINITIVES

#### 2.1 Introduction

Beginning with Lees (1960), a considerable literature has emerged on a linguistic entity referred to as a "nominalization" or "derived nominal". In general terms, these are nouns or noun phrases that seem to be derived from lexical entries that belong to categories other than noun. In some instances they keep some of the characteristics of the categories that they come from; in others, they do not. Here, we are interested in those nominals that derive from a verbal counterpart.

The first approach to this phenomenon is represented by Lees (1960) (and others later), who argues that these structures are derived from sentences. This approach is called "transformational". A second approach has surfaced, initiated by Chomsky (1970), called "lexicalist," that counters Lees' claims. The options have grown over the years, but this main distinction stays at the heart of the debate.

Section 2.3 is devoted to the problem of the derivation of nominalizations. The question asked in that section is the one concerning the locus of the category change. Does the change in category from verb to noun occur in the lexicon or in the syntax, or in both?

### 2.1.1 Properties of Nominalizations

The properties of nominalizations can be summarized in the following way:

#### **General Properties**

- They may have the internal structure of a NP.
- They may have the internal structure of a VP.
- The subject is usually optional.
- In NP-like nominalizations, if the object is present, the subject is not optional.
- Word order may differ from VP word order.
- Depending on the type of nominalization, restrictions apply on the Determiner system.
- Either adverbs or adjectives may modify the head of the nominalization, depending on the type.

#### **Language-specific Properties**

- In English, in the case of the NP-like nominalizations, the subject is in the Genitive case and the object can only appear preceded by "of".
- In Basque, restrictions on inflection are different from the ones present in an inflected VP.

### 2.1.1.1 Distribution

Nominalizations may occupy all the positions occupied by a regular NP, but also exhibit, in the case of -ings or infinitivals, the distribution of a VP. I will take examples from Pullum (1991):

The external syntax of a NP.

The subject cannot be extracted.

- (1) a. *You were upset at whose being too lazy to accompany you?*  
b. *\*Whose were you upset at [NP e] being too lazy to accompany you?*
- It occurs in S-internal NP positions, unlike subordinate clauses.
- (2) a. *That the world is round must be obvious for birds.*  
b. *\*Must that the world is round be obvious for birds?*  
c. *My leaving now would be taken as an admission of defeat.*  
d. *Would my leaving now be taken as an admission of defeat?*

what he calls the lexicalist approach. He argues (1) that the relationship between nouns and the sentences from which they derive is non-productive:

(7) a. *John is easy to please*  
b. \**John's easiness to please*

(2) that the meaning relationship between the derived noun and the verb is not predictable:

*perform / performance*

*delegate / delegation*

and (3) that derived nouns share properties with non-derived nouns. These properties are:

- a. they can follow determiners
- b. they can follow adjectives
- c. they cannot be modified by auxiliaries
- d. they cannot be modified by modals
- e. they cannot occur with adverbs

(8) a. *the performance*  
b. *the beautiful performance*  
c. \**the having performance*  
d. \**the most performance*  
e. \**the performance beautifully*

Chomsky compares gerundive nominals and derived nominals and notes that, while gerundive nominals can be formed quite freely from verbs, derived nominal formation is much more restricted. He assumes that the latter are introduced by rules of the categorial components of the base, while gerundive nominals are formed through general rules of affix placement. Derived nominals are, thus, essentially unpredictable, while gerundive nominals, even though restricted, are predictable.

#### 2.1.1.3 Control

The concept of control has many definitions in linguistic theory. In its broadest sense it is concerned with any non-freely interpreted empty category.

One of the characteristics of clauses or noun phrases headed by nominalized verbs is the absence of some of the arguments, most often the subject. Therefore, there is an important interaction between a theory of control and an account of nominalizations.

argues that control infinitivals are IPs unless otherwise required by s- or l-selection applied from the higher predicate, the case requirements of the infinitival subject, and the ECP.

In order to account for the data in (11), Bošković adopts the idea that [+Tense] non-finite INFL can check null Case. Control infinitives, in contrast to ECM infinitives, are marked for Tense. Thus, PRO is Case-checked under SHA (Spec-Head Agreement) with [+Tense] non-finite INFL.

(11)a. *John believed [AgrP him; t; [IP t; to be crazy]]*  
b. \**John believed [IP PRO to be crazy]*  
c. *John tried [CP PRO to win]*

In some instances PRO infinitives must be IPs, in others, such as the infinitival complement of *croire*-class verbs, in French, they must be CPs.

Johnson (1988) argues that subjects of gerunds are Case-marked by an external governor. This element can be a preposition:

(12) *Liz left without you telling a story*

He argues that gerunds can be S or S'. If they are S the subject can be case-marked by an external governor. If they are S', the alternation between NP and PRO is available:

(13) *I remember [S' [S PRO telling the story]]*.

Zabala (1995) argues that in “infinitival” clauses in Basque the empty subject must be pro rather than PRO. Her arguments are based on the fact that these empty categories are not linked to any other argument, they don’t need to be controlled or have arbitrary reference, and they can appear in argument positions other than subject. We will explore these issues in relation to Basque in the next chapter.

The notion of control is directly linked to the correct understanding of nominalizations. As seen above, if we adopt a configurational approach, the interpretation of control phenomena allows us to decide what type of phrase we are dealing with. From a functional or a semantic point of view, the analysis of control phenomena allows us to categorize nominalizations both semantically and functionally.

### 2.1.2 Computational Problem

Within Natural Language Processing, hereafter NLP, verbal nominalizations are one of the most interesting linguistic phenomena. Very few instances of this phenomenon can be accounted for through electronic dictionaries, and for that reason, nominalized structures present a serious computational challenge in any language.

Nominalizations are mutants i.e., entities that have undergone a categorial change without completely becoming an entity belonging to the new category and, as such, difficult to capture. Outside of the field of Computational Linguistics, in the context of Theoretical Linguistics, there is a good deal of debate over these structures. Computational Linguistics, dealing with computer applications, tries to inherit the results of work done in Theoretical Linguistics. In what concerns nominalizations, however, Theoretical Linguistics offers little help.

Derived nominals, even though mostly nominal, also exhibit verbal behavior as to, for example, their complements. This linguistic entity (nominalization in general) takes different shapes depending on the language. In English we see derived nominals, either through zero-derivation (*fear*) or through affixation (*arrival*), and -ing nominalizations (*understanding*). Infinitivals in English can also be considered nominalizations (*to be good is to be great*). In the Romance languages we see derived nominals similar to the English ones (*arrivée* 'arrival' in French, *construcción* 'construction, building' in Spanish, etc.) and nominalized infinitives. Nominalized infinitives seem to display very different behavior from language to language. In Basque, we have derived nominals (*eros-keta* (from *erosi*, 'to buy'), 'purchase'), -*te*/nominalizations (*eros-te-a* (from *erosi*, 'to buy'), 'buying, purchase'),

and nominalized infinitives (*etorri-a* (from *etorri*, 'to come'), 'arrival, coming').

In the following sections, I will discuss the properties and distribution of these structures.

### 2.1.3 Analysis

The most important difficulty one encounters when dealing with nominalizations from a syntactic point of view concerns the issue of endocentricity. For any X-Bar theory of syntax endocentricity is a required property. This property states that the category of a phrase and that of the head of that phrase have to be the same.

Nominalizations do not seem to respect endocentricity, since they belong to nominal phrases (NP or DP), but their heads are verbal, as we saw in section 2.1.1.1. Authors have tried to circumvent the problem in different ways. Some have posited two different projections (Grimshaw, 1990) to avoid the violation of endocentricity. Others have posited mixed projections (Bresnan, 1997), leaving unresolved the endocentricity issue. Others have changed the characterization of the phrases in question, DPs (Abney, 1987). We will get into the details of these analyses later on.

For other authors who work in the context of syntactic frameworks that posit a less strict notion of endocentricity (Pullum, 1991), the

classes of nominalizations. The first class includes “action or state nouns”. Action or state nouns come from verbs, and describe the process or the event of the action. This type also has other readings, but what differentiates it from the other class of nominalizations is the fact that it retains some of the characteristics of the verb or adjective it derives from. Examples of the second class are ‘creation,’ ‘arrival,’ ‘truck-driving’. These nominalizations behave syntactically like other nouns in the language, while they only keep (often unpredictable) semantic and morphological relations to the associated verb or adjective.

Nominalizations are further divided into agentive nominalization (*singer*), instrumental nominalization (*slice*), manner nominalization (*his walking* (both interpretations: manner and fact), objective nominalization (*umcabango* ‘thought’, in Zulu, from *cabanga* ‘think’), locative nominalization (*mulató* ‘place for having fun; bar’, from *mulat* ‘to have fun’, in Hungarian) and reason nominalization (*paJindit* ‘reason for leaving’ from *indit* ‘to leave’, in Sundanese). Their conclusions are that processes for forming nouns are likely to be irregular and unpredictable. In addition, they note that languages vary in the degree to which these nominalizations will resemble regular NPs or the sentences from which they “derive”.

Koptjevskaja-Tamm (1993), in her typology of nominalizations, seems to come to similar conclusions.

In the following sections I will discuss more closely different types of nominalizations in English and in three Romance languages. This will provide the reader with a better idea of the range of possibilities and the issues involved.

### 2.2.1 English

As mentioned above, we need to make a distinction between nominals and -ings in English. Even though nominals formed through “-ing derivation” behave, in some respects, like other derived nominals, this formation exhibits a set of behaviors that are not found in other nominals.

Section 2.2.1.1 will be devoted to nominal -ing formations, while in section 2.2.1.2, I will discuss other derived nominals in English.

#### 2.2.1.1 -ings

Also called gerunds, these constructions in English exhibit a set of behaviors that go from that of a VP to that of NP. Pullum (1991) and Abney (1987) provide analyses of these structures in English from two very different syntactic perspectives.

not problematic. This mechanism allows him to posit what he calls “heterocategorial heads”.

Through the FCR above, he allows for the category of the head to be verbal, while the category of the phrase is nominal.

There are several problems with this account. First, the role of this suffix ‘ing’ appears to be totally accidental. Second, it remains unclear how nominal behavior spreads internally into a VP headed by a noun. These categories appear to be more “mixed” than “heterocategorial”.

Bresnan (1997) argues for the existence of mixed categories, which she defines as: *a single word heads a phrase which is a syntactic hybrid of two different category types* (p.2). Within the Lexical Functional Grammar framework, she posits the Extended Head Theory, by which different categories share the same head in f-structure, not c-structure. In this manner, she accounts for the mixed behavior of nominalizations.

She argues that positing a lexically indeterminate head would not account for the fact that in several languages the category of the phrase is clearly nominal. In addition, these structures exhibit phrasal coherence constraints, i.e. the order of the elements inside the phrase is fixed. As Zucchi (1993) shows, in the nominalized infinitive in Italian the constituents preceding the infinitive are always nominal, while the ones

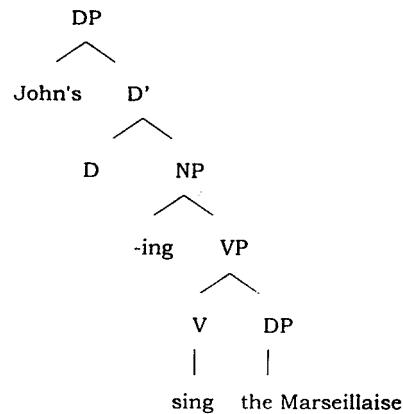
following the infinitive are verbal. The choice has to be made between VP constituents and NP constituents.

LFG is incompatible with morphological derivation in syntax because of its lexical integrity principle (see Sells, 1995) and because of the absence of serial derivational mechanisms from its architecture.

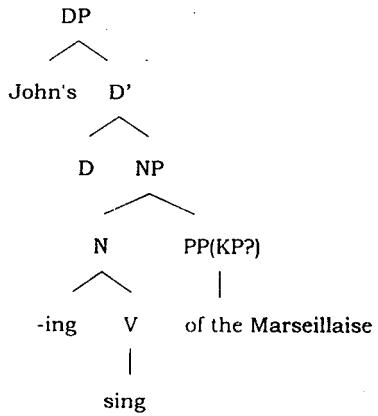
c-structure represents overt forms of expression, while f-structure represents equivalence classes of c-structure. The f-structures of the N and the VP can be then identified through unification and share attributes. A suitable theory of argument structure is still needed to account for the selection problems, i.e., how do we predict that in different languages different nominalized structures (nominalizations or infinitivals) can be of mixed category? In Italian, infinitivals can head mixed categories, while nominalizations cannot. In other languages, such as Catalan, the opposite is true.

In the next section we discuss Abney’s (1987) approach which relies on the parallelism the author establishes between NPs and sentences.

b. *Poss-ing*: *John's singing the Marseillaise*.



c. *Ing-of*: *John's singing of the Marseillaise*.



Abney's DP hypothesis has the advantage of providing a uniform account of the behavior of -ing nominalizations in English. It also gives functional elements a uniform structural position in nominal and verbal systems.

In the next section we will discuss nominalizations not formed through -ing affixation in English.

#### 2.2.1.2 Other derived nominals

English presents the following types of deverbal derived nominals:

Action/state nouns *destroy* → *destruction*

*arrive* → *arrival*

*fear* → *fear*

*rob* → *robbery*

*behave* → *behavior*

Agentive nouns *sing* → *singer*

Instrumental nouns *slice* → *slicer*

Object nominalizations *paint* → *painting*

As we saw in the preceding section, tests such as the one in (21) (from Grimshaw 1990) allow us to distinguish between those readings in cases of nominalizations that are ambiguous between the two readings, such as 'examination'.

In the next section, we discuss the notion of argument structure in nominalizations, developed by Grimshaw (1990).

#### 2.2.1.2.2 Complex Event Nominals and Argument Structure

Grimshaw (1990) makes a three-way distinction among result, simple event, and complex event nominals. She argues that only complex event nominals and -ings have an argument structure. She considers -ings to be simply verbal and to have argument structure, while complex event nominals inherit their argument structure from the verb from which they come. Complex event nominals and related simple event and result nominals have a related lexical conceptual structure, but only complex event nominals have an event structure and a syntactic argument structure.

She groups together simple event nominals, such as 'trip' and result nominals, such as 'exam'. They can be preceded by demonstratives, as result nominals do. They occur with optional modifiers, rather than arguments. They pluralize. They do not allow

adjectives such as 'constant,' unless they are in the plural, and they do not manifest event control:

- (22)a. *That trip took three days*
- b. *Yesterday's trip was a success*
- c. *Those trips were nice*
- d. *\*The constant trip was boring*
- e. *The constant trips became a problem*
- f. *\*That trip in order to...*

Result nominals take only modifiers. Simple event nominals take complements and modifiers, but do not take arguments. Complex event nominals take arguments. Complements are licensed by the lcs (lexical conceptual structure) of the base verb, but not by its a-structure (argument structure). What is predicted here is that if we say that complex event nominals project an argument structure like verbs do, the complements to complex event nominals will be obligatory. (*The felling \*(of the trees)* and *They felled \*(the trees)*)

Grimshaw makes very little mention of -ings. She views them as being ambiguously verbal or nominal. The category of the head can only be discovered through certain tests. In general, she assumes they are verbal, and because they are verbal they theta-mark like any other verb and unlike result or simple event nominals.

### 2.2.2.1 Catalan

Fullana (1995) argues that Catalan nominal infinitives have only a manner reading and provides an explanation for that fact.

(24) *El parlar d'en Joan*

the talk/N of the Joan

The example above is ambiguous between an object reading (Joan's idiolect) and a manner reading ('Joan's way of talking'). She assumes that the object reading is a lexicalized one since it is neither regularly created nor regularly interpreted.

Relying on a Lexical Semantics framework, she postulates that nominal infinitives with a manner reading are due to two lexical rules:

- a. Rules of semantic composition that subordinate the Lexical Conceptual Structure of a verbal root to an abstract value that we call manner giving rise to a complex LCS where manner is the head.
- b. Morphological rules that add a morpheme *-r* to the verbal root which is responsible for the nominalization of the verb.

Morphological Process:

$[Xv] \rightarrow [ [Xv] -r N ]$

### Semantic Process

LCS1: [Xv] LCS2: [MANNER Y X]

LCS: { MANNER [X V] X }

As other authors have assumed that *-ing* is a nominalizing affix, she assumes that *-r* also has this function in Catalan.

The complement of a nominal infinitive that expresses manner is always obligatory. Since MANNER is the head of the construction, the complement is obligatory inasmuch as it involves inalienable possession.

Only a few adjectives can appear in phrases with the manner reading. They are the same adjectives that can qualify the noun *manera*: *lent* 'slow,' *dolç* 'sweet,' etc.

#### Nominal infinitives with a manner reading:

- Cannot act as an argument of certain predicates

(25) *\*lamento el caminar d'en Pere*

regret the walk/N of the Pere

I regret Pere's (way of) walking'

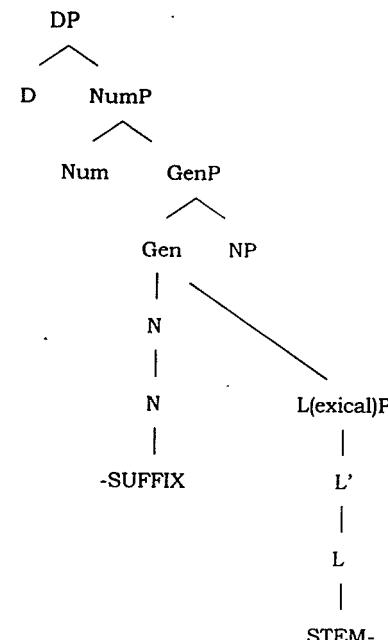
She distinguishes between lexical and syntactic nominals and characterizes them as follows:

Syntactic nominals:

- The DP will receive the event/process interpretation.
- The nominalization affix-heads a functional projection NP that takes as a complement a category-neutral lexical projection headed by a stem. This L<sub>0</sub> "becomes" an NP at the syntactic component.

(28) *El repartiment de queviures*

'The distribution of food'



She characterizes lexical nominals as follows:

- The nominalization affix is a derivational morpheme.
- These are result nominals.
- Suffixation is implemented at D-structure.

(This applies also to non-deverbal nominals)

g. *El descobriment de Koch del bacil de la tuberculosi*

'The discovery by Koch of the tuberculosis bacillus'

In (31a) and (31b) we see how a nominalization that is ambiguous between an event and a result reading gets disambiguated through passivization, by making the agent explicit. (31b) is, thus, an event nominalization, while (31a) has a result reading. In (31c) and (31d) we see two different affixation processes used to derive nominals in Catalan, (31c) being an event nominal and (31d) a result. (31e) and (31f) show a contrast that demonstrates that (31c) is an event nominal, through passivization in (31f) and through the restriction of having two 'de' modifiers, one being the subject and the other the object. (31g) shows that (31d) is indeed a result nominal, since it can have the Agent argument in the genitive followed by the object argument.

We have thus seen that in Catalan nominalized infinitives can only have a manner reading and seem to be derived morphologically. Also, while presenting another approach to the analysis of nominalizations, i.e. by positing a category-neutral stem and two loci of operation (the syntax and D-structure), derived nominals in Catalan seem to allow result and event nominalizations. We also saw that there are nominalizing affixes in Catalan which are specialized to derive event nominalizations and others that derive result nominalization.

In the next section we discuss the same type of facts in Italian.

2.2.2.2 Italian

Zucchi (1993) looks at infinitivals in Italian and argues that there are three types of infinitivals in Italian:

- N-infinitival NPs

(32) *Il mormorare somesso del mare*  
the murmur-inf soft of the sea  
'The soft murmuring of the sea'

- S-infinitival NPs

(33) *L'avere egli compiuto i primi studi in*  
the his/her have-inf complete-pple the first study-pl in  
*Francia*  
France

'The having he done his first studies in France'

- VP-infinitival NPs

(34)a. *Il suo mormorare parole dolci*  
the his/her murmur-inf word-pl sweet  
'His soft murmuring sweet words'

b. *Il suo mormorare somessamente*  
the his/her murmur-inf softly  
'His softly murmuring'

optional and the distribution does not correspond to the distinction between NPs and VPs.

- (36)a. *Que vengas mañana me hace mucha ilusión*  
that come/2sg.subj tomorrow me-D do/3sg.pres.ind much pleasure
- b. *Me hace mucha ilusión que vengas mañana*  
'Your coming tomorrow makes me very happy'
- c. *El que vengas mañana me hace mucha ilusión*  
the come/2sg.subj tomorrow me-D do/3sg.pres.ind much pleasure
- d. *Me hace mucha ilusión el que vengas mañana*  
'Your coming tomorrow makes me very happy'
- e. *Es imposible comprender sus razones*  
be/3sg.pres.ind impossible understand-inf his/her reasons  
'It is impossible to understand his/her motivation'
- f. *\*Es imposible el comprender sus razones*
- g. *Saber no ocupa lugar*  
know-inf not occupy/3sg.pres.ind space  
'To know does not take room'
- h. *El saber no ocupa lugar*  
the know-inf not occupy/3sg.pres.ind space  
'The knowing (knowledge) does not take room'

- i. *Saber mucho no ocupa lugar*  
know-inf a lot not occupy/3sg.pres.ind space  
'To know a lot does not take room'

- j. *\*El saber mucho es bueno*

As we see in (36a), (36b), (36c), and (36d), the presence of the determiner may be optional with a tensed clause<sup>3</sup>. In (36e), (36f), (36g), (36h), (36i), and (36j) we see that the determiner is not optional with an infinitival in certain environments. The contrast between (36h) and (36i) shows that when the infinitival appears with an object, the determiner is no longer possible. Example (36f) shows that in an environment in which a clause is expected the determiner is not possible either. Thus, Plann's observation is correct but needs to be refined. When a nominalized infinitive in Spanish has a determiner it is either because the environment requires a NP or because either an NP or a clause is acceptable. In such cases the presence of the determiner is optional. In environments where a clause is required, the nominalized infinitive can not have a determiner

Plann also points to the fact that there is a semantic contrast between the uses with and without the determiner.

- *el* + S may conjoin with other NPs, but bare S never does:

(38)a. *Siempre evito las duchas frías y \*(el) hacer ejercicio*  
 always avoid/1sg.pres.ind the cold showers and (the)  
 exercise-inf  
 'I always avoid cold showers and exercising'

In my dialect I obtain the exact opposite judgment. The example is acceptable only without the determiner.

Yoon and Bonet-Farran (1991) distinguish DP (Determiner Phrase) infinitives from CP (Comp Phrase) infinitives, the former being nominal while the latter are sentential.

Among the DP infinitives they distinguish three types:

(39)a. *[El cantar yo la Traviata] traerá malas consecuencias*  
 the sing-inf I the Traviata bring/3sg.fut.ind bad  
 consequences  
 'My singing the Traviata will bring bad consequences'

b. *[El tocar de la guitarra de María] es muy elegante*  
 the play-inf of the guitar of María be/3sg.pres.ind very  
 elegant  
 'Maria's guitar playing is very elegant'

- c. *[El tocar la guitarra de Juan] me pone nerviosa*

the play-inf the guitar of Juan me-dat get/3sg.pres.ind  
 nervous/fem.sg  
 'Juan's playing the guitar makes me nervous'

They explain that (39a) is sentential in all but its external distribution, (39b) is nominal both internally and externally, and (39c) is nominal with respect to material outside VP but is verbal inside VP, shown by the accusative case in its object. The three types correspond to the English Acc-ing, Ing-of, and Poss-ing, respectively. Example (39b) while marked as acceptable by Yoon and Bonet-Farran strikes me as totally ungrammatical<sup>3</sup>. And it cannot really be argued that their analysis is valid for Spanish speakers who accept (39b) either because the English counterparts do not show the same variation as (39b) do in Spanish. I do not believe that the parallelism that the authors are trying to build with Abney's analysis of English -ings is valid. Only two types of infinitival DPs exist in Spanish: Acc-ing and Poss-ing. As we will see in more detail later on, not only do passive nominalized structures in Spanish present a restriction by which the Agent of a nominalization in which the object was introduced by the preposition 'de' can only be introduced by 'por parte de', but also, infinitivals do not allow objects to be introduced by 'de.' Infinitivals allow the Agent to be introduced by 'de,' though.

It seems to me that we have an issue of emphasis here. The subject of the infinitival is not controlled, thus, when this subject needs to be specified we have two possibilities in Spanish: we either place it postverbally if it is a pronoun; or use a tensed clause:

(44)a. *Cantar yo ahora sería una estupidez*

    sing-inf I now be/3sg.cond a stupidity

    'For me to sing now would be stupid'

b. *Que María cante ahora sería una estupidez*

    that María sing/3sg.pres.subj now be/3sg.cond a stupidity

    'For María to sing now would be stupid'

c. *Que cante María ahora sería una estupidez*

    that sing/3sg.pres.subj María be/3sg.cond a stupidity

    'For María to sing now would be stupid'

I do not have an explanation for these facts, but they certainly contradict the analysis of AGR proposed by Yoon and Bonet-Farran.

The presence of the subject is allowed when the infinitival is introduced by a preposition.

(45)a. *Ana salió de la habitación /sin verla yo/ sin yo verla/*

    Ana leave/3sg.pret.ind of the room without see-inf her I/

    without I see-inf her

    'Ana left the room without me seeing her'

b. *Ana salió de la habitación /sin verla Juan/ sin Juan verla/*

    Ana leave/3sg.pret.ind of the room without see-inf her  
    Juan/ without Juan see-inf her

    'Ana left the room without Juan seeing her'

In this case, the preposition governs IP and case can be assigned to the subject.

De Miguel (1995) argues that there are two types of Spanish infinitivals headed by a determiner, rather than it being a mixed structure. One of them is inflectional in nature and projects clausal structures, while the other is derivational and projects nominal structures.

She argues that in inflectional infinitivals the subject is in the nominative, while it is "genitive" in the derivational structures, this contrast is shown in (46a) and (46b). In addition, in the first type of infinitivals the subject cannot be realized as a genitive (46c):

(46)a. *El andar el niño por ahí nos preocupa.*

    the walk-inf the boy around us-D worry/3sg.pres.ind

    'The boy's walking around worries us'

b. *El andar errabundo del niño nos preocupa.*

    the walk-inf aimless of the boy us-dat worry/3sg.pres.ind

    'The aimless walking of the boy worries us'

b. \**El dulce haber hablado de Juan...*  
 the sweet have-inf talk-pplc of Juan..

Derivational infinitivals may be replaced by masculine pronouns (52b), while inflectional infinitivals can only have neutral pronouns (52a):

(52)a. *El hablar Juan con ella era tan importante que ya no pudo vivir sin ello.*  
 the talk-inf Juan with her be/3sg.imp.ind so important that no more not can/3sg.pret.ind live-inf without it  
 'Juan's talking to her was so important that s/he could not live without it anymore'

b. *El dulce hablar de María era tan importante que ya no pudo vivir sin él.*  
 the sweet talk-inf of María be/3sg.imp.ind so important that no more not can/3sg.pret.ind live-inf without it  
 'María's sweet talking was so important that s/he could not live without it anymore'

Inflectional infinitivals cannot have restrictive relative clauses (53a), while derivational infinitives do (53b):

(53)a. \**el hablar Juan que oyó ayer...*  
 the talk-inf Juan that hear/1sg.pret.ind yesterday..

b. *El hablar de Juan que tanto me gusta...*  
 the talk-inf of Juan that so much me-D please/3sg.pres.ind  
 'Juan's talking that I like so much..'

Inflectional infinitives accept only non-restricted relatives headed by the neuter forms *lo cual*, *que*, *lo que* (54a) and (54b), while derivational infinitives allow only *el cual* or *que* (54c) and (54d):

(54)a. *El hablar Juan, lo cual tuvo lugar ayer.*  
 the talk-inf Juan, which have/3sg.pret.ind place yesterday  
 'Juan talking, which took place yesterday..'

b. \**El hablar Juan, el cual tuvo lugar ayer*  
 the talk Juan, which have/3sg.pret.ind place yesterday

c. \**El dulce hablar de Juan, lo cual tuvo lugar ayer.*  
 the sweet talk-inf of Juan, which have/3sg.pret.ind yesterday

d. *El dulce hablar de Juan, el cual tuvo lugar ayer*  
 the sweet talk-inf of Juan, which have/3sg.pret.ind yesterday  
 'Juan's sweet talking, which took place yesterday'

Only infinitivals of the same type may coordinate:

(55) \**Ese dulce hablar de Juan y el cantar Elena.*  
 that sweet talk-inf of Juan and the sing-inf Elena..

### 2.3 Derivation: Lexical versus Syntactic

Various attempts have been made to characterize the interaction between syntax and morphology in the context of these category changes.

Yoon (1996) argues that the -ing affix in NGPs in English does not function as a nominalizing suffix. The author proposes that both lexical and phrasal nominalizations are the result of zero-derivation from projections of the present participle form of the verb, applying respectively to  $V^0$  and  $V^n$ . This analysis presupposes a view of morphosyntactic interaction in which morphological operations such as affixation and zero-derivation apply in both the lexicon and syntax.

Pullum (1991) notes that there is a recurring cross-linguistic generalization. The same (de) verbal form is used in both lexical and phrasal nominalizations, not only in English but in several languages.

The infinitival in Spanish nominalization behaves like the NGP. It takes verbal clitics (56a) and it does not pluralize (56c), properties not found in lexical nominalizations.

(56)a. *El partíles la cara todos los días no está bien*

the break-inf they/3pl.D the face every day not  
be/3sg.pres.ind well  
'It's not right to smack them every day'

b. *El ir mucho al cine es una buena costumbre*

the go-inf much to the movies be/3sg.pres.ind a good habit  
'To go often to the movies is a good habit'

c. *\*Los ires mucho al cine son....*

the go-inf-pl much to the movies be/3pl.pres.ind...

For Yoon (1996) the infinitival suffix in Spanish does not function as a nominalizer; it is the verb stem plus the infinitive suffix that is zero-derived to form a noun.

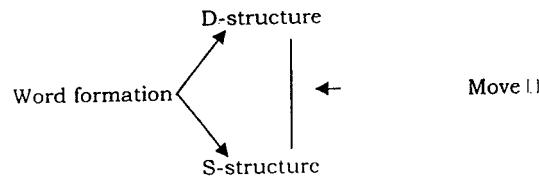
He assumes three types of affixes:

- (i) Affixes that are strictly lexical
- (ii) Affixes that can be both lexical and syntactic
- (iii) Affixes that are always syntactic

Affixes do not have a BAR level specified for them. If they apply at a BAR level 0 ( $n=0$ ), the result is a lexical derivation. If they apply at a level 2 ( $n=2$ ), the derivation is syntactic. Then, the zero-affix involved in nominalizations in Spanish is of type ii, whereby it applies to BAR levels

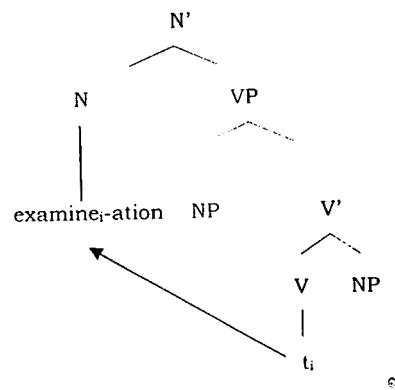
$$0 \leq n \leq 2.$$

Fu (1994), following Borer's Parallel Morphology model, tries to account for derived nominals in Chinese. In Parallel Morphology, the word formation component is autonomous and interacts with both the lexicon and the syntax:



When the interaction takes place in D-structure, the properties of the non-head elements inside the phrase are opaque; when word formation applies at S-structure, they are transparent.

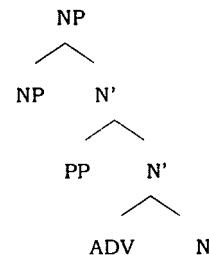
He posits V-to-N raising. N selects VP, and since it is selected, it is not a barrier for movement.



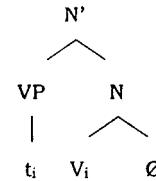
72

This structure allows the author to account for some VP-like behavior presented by derived nominals. Using Parallel Morphology, two structures are projected, one resulting from a D-structure derivation (*trip, race*), and the other from a post-D-structure derivation (action nominals, V-ing of). While lexicalist approaches predict that the derived nominal will have the structure of a NP, Parallel Morphology predicts that it will have structural similarities to a VP and will, indeed, contain a VP.

Lexicalist representation:



Parallel Morphology:



73

constitute separate syntactic entities at D- and S-structure. So, for her there are two different types of suffixes.

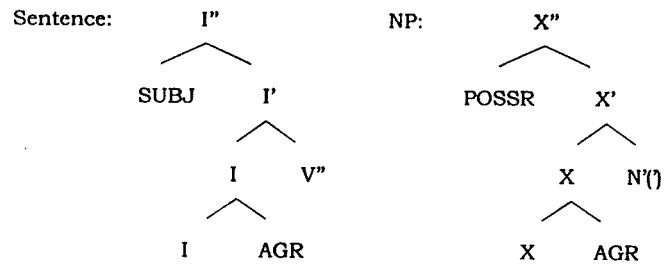
The nominalization affix heads a functional projection NP that takes as a complement a category-neutral lexical projection headed by a stem. This L<sub>0</sub> "becomes" an NP at the syntactic component.

Abney (1987) uses facts from Hungarian, Turkish, Yupik and Mayan to introduce the idea that an inflectional head is present in the NP.

Hungarian data (taken from Szabolcsi: 1987):

- (59)a. *az en kalap-od*  
the I:NOM hat:1st sg  
'my hat'
- b. *a te kalap-od*  
the you:NOM hat:2sg  
'your hat'
- c. *a Peter kalap-ja*  
the Peter:NOM hat:3rd sg  
'Peter's hat'

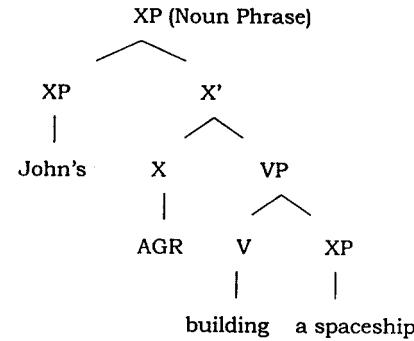
Thus, Abney posits:



Turkish data (from Underhill: 1976):

- (60) *Halil'-in kedi-ye yemek-0 ver-me-dig-i*  
Halil:GEN cat:DAT food: ACC give:NEG-ING-3sg  
'Halil's not giving food to the cat'

Thus, he proposes:



the ones that are needed to justify, for example, the theta role assignment capabilities of nominalizations. Finally, it can be extended to other languages to produce the desired results.

#### 2.4 Conclusion

In this chapter, I have presented general, as well as language-specific, properties of nominalizations. I have discussed linguistic issues that are brought up by nominalizations such as control phenomena or derivational versus inflectional morphology.

In addition, through data from English and three Romance languages, I have presented not only the range of manifestations of the phenomenon of nominalization, but also the different analyses proposed in the literature for such facts.

### CHAPTER 3

## **NOMINALIZATIONS IN BASQUE**

#### 3.1 Introduction

I start this section by providing the reader with a brief description of general properties of Basque that will be useful in understanding both the data and the issues presented later.

Basque has been classified as an ergative language with free word order and no prepositions. Determiners and postpositions are attached at the end of the phrase. The verb usually appears at the end of the Verb Phrase. Most numerals and quantifiers appear at the beginning of the phrase (Saltarelli 1988, Laka 1996).

Verbal inflection in Basque is usually formed by the root or participial form of the verb and an auxiliary verb. This auxiliary verb takes the tense as well as the markers for the three participants of the sentence: the subject, the direct object, and the indirect object. The verbal form also carries number markers for the participants. In some

Section 3.2 is devoted to a typology of nominalizations in Basque. Section 3.3 constitutes the bulk of the chapter in that the core issues of the problems posed by nominalizations are discussed there. I will present previous accounts and argue for an approach to Basque nominalizations along the lines of Grimshaw (1990). In section 3.4, I discuss the derivation of nominalizations in Basque.

As was mentioned in Chapter 1, nominalizations are viewed in this study as a theoretical linguistics problem, a computational problem, and a tool of diagnosing language attrition from an Anthropological Linguistics point of view. Section 3.5 offers a partial computational account for Basque nominalizations. The Anthropological Linguistics approach and fieldwork results will be discussed in Chapter 4 and Chapter 5.

### 3.1.1 Ergativity

Basque is considered an ergative language, at least in its morphology. But morphological ergativity does not entail syntactic ergativity<sup>6</sup>. This is to say that, frequently, syntactic processes that make reference to the notion of subject and object seem to apply in ergative languages, rather than the notions 'ergative' and 'absolutive'. A certain amount of debate exists on the question of Basque being syntactically ergative. Several pieces of evidence have been put forward to argue for

Basque not being syntactically ergative. Interestingly enough, most of the data provided involves nominalized structures and infinitivals.

Levin (1983) presents some control structures in Basque:

(64)a. *Gaur-ko kazeta eros-te-ko eskatu diot*  
today-G newspaper-A buy-nom-G ask-prf  
Aux/3sgA.3sgD.1sgE (*ukan*)  
'I asked him to buy today's newspaper'  
b. *Berehala joa-te-ko esan dio*  
Immediately go-nom-G say-perf Aux/3sgA.3sgD.3sgE  
(*ukan*)  
'S/he told him/her to go immediately'

In (64a) the missing argument is the transitive subject; and in (64b) the intransitive subject is missing, but the object of the transitive verb could not be missing. She argues, thus, that subjects group together.

Ortiz de Urbina (1989) offers additional evidence for proving that Basque is not a syntactically ergative language:

b. *Semea eskolan utzi zuen eta klasera joan zen*  
 son-A school-at leave Aux/past.3sgA.3sgE (*ukan*) and  
*klasera joan zen*  
 class-adl go Aux/past.3sgA (*izan*)  
 '(S/he) left his/her son at school and went to class'

In (68a) the auxiliary for the first verb, *utzi*, has been dropped, while in (68b) it is present. Note that the auxiliary for the first verb is *ukan* 'have' and *izan* 'be' for the second verb. In principle, it should not be possible to coordinate two structures where one of the subjects is marked with the ergative and the other with the absolute if "subject" was not a relevant notion, assuming that coordination requires identity as to case marking.

In partitive constructions in Basque, objects of transitive verbs and subjects of intransitive verbs pattern together as opposed to subjects of transitive verbs. But Ortiz de Urbina (1989) argues against this as proof of syntactic ergativity on the basis that this can be accounted for by means of the Unaccusative Hypothesis.

The Unaccusative Hypothesis was first put forward by Perlmutter (1978) and adapted to the Government and Binding (hereafter GB) framework for syntactic theory by Burzio (1981). This hypothesis splits intransitive verbs into unaccusative and unergative verbs. Unaccusative

verbs have a single D-structure argument, i.e. the subject of this type of verbs is best analysed as an underlying object. Unergative verbs have a D-structure subject. The relevant opposition is traditionally represented by the following Italian examples:

(69)a. *Arrivano molti studenti*  
 'Many students arrived'  
 b. *Ne arrivano molti*  
 'Many of them arrived'  
 c. *Telefonano molti studenti*  
 'Many students telephoned'  
 d. *\*Ne telefonano molti*  
 'Many of them telephoned'

Assuming that *arrivare* is an unaccusative verb in Italian that assigns a θ-role to a single D-structure argument (the object), and *telefonare* is an unergative intransitive with a D-structure subject argument, *ne*-cliticization is restricted to deep objects.

The distinction, even though relevant to the syntax, depends on semantic arguments rather than morphological case marking. Levin (1983) proposes that in Basque all verbs that take *izan* as an auxiliary are to be considered unaccusative. Ortiz de Urbina (1989) notes, though, that in Basque, there is not a one-to-one correspondence between

The distinction between a D-subject (Deep Structure subject) and a D-object (Deep structure object), proposed in the Unaccusative Hypothesis, can explain the distribution of the partitive in Basque. Only absolute arguments can appear in the partitive. By the Unaccusative Hypothesis, both subjects of intransitive unaccusative verbs and objects of transitive verbs are actually D-objects. Then the contrast can still be explained in terms of a subject/object contrast. This argues for an analysis in which Basque is not syntactically ergative. Rather it seems to present a case of the 'extended ergative' type as in Dixon (1972), whereby the language is syntactically accusative but shows morphological marking on the subject, while the object is morphologically unmarked.

Laka (1996) posits that in partitive structures the partitive is a polar determiner, like *any* in English. For her, then, the behavior of the partitive has nothing to do with absolute or ergative

Rebuschi (1989) agrees that Basque is not syntactically ergative, but raises the question of whether this is a consequence of a configurational organization of the Phrase Structure, or if it is a logically independent phenomenon. He explores the idea that Basque does not really have a VP as a constituent. The absence of a VP allows the verb to govern its subject.

In Arrieta, Joseph & Smirniotopoulos (1986) some arguments are put forward to show that Basque may well be syntactically ergative. They put forward examples such as:

(72)	<i>Bata</i>	<i>joan</i>	<i>da</i>	<i>besteak</i>
	one-Asg	go-perf	Aux/3sgA(izan)	other/Apl or Esg
	<i>ikusi</i>	<i>gabe</i>		
		see-inf	without	
			'One went without the others seeing him/her or	
			without seeing the others.	

As (72) shows, in certain environments, the missing argument can be the subject (ergative in the subordinate and absolute in the main clause) or the object of the subordinate clause (absolute subject in the main clause and absolute object in the subordinate clause). See Arrieta, Joseph, and Smirniotopoulos (1986) for further evidence.

For the purpose of this study it is not crucial to decide where the truth lies. The issue remains open. Still, the preceding observations clarify the basics of Basque structure.

In the next section I discuss another open issue concerning Basque: configurationality.

There are also restrictions at the phrasal level. Noun complements and relative clauses occur to the left of the head while adjectives occur to the right of the head noun.

- (77)a. *Etxearen erosketa*  
house-A-G buy-nom  
'The purchase of the house'
- b. *Etxe haundia*  
House big-A  
'The big house'
- c. *Etxeko andrea*  
house-of woman-A  
'The housewife'
- d. *Abestu dun gizona*  
sing-prf Aux/3sgE.3sgA-comp(ukan) man-A  
'The man who sang'

The above list of examples is not exhaustive and is only meant as a sample of the restrictions that Basque imposes on word order.

Ortiz de Urbina (1989) argues that Basque is a configurational language. He puts forward data concerning anaphors:

- (78)a. *Bere; amak Jon; atzo deitu zuen*  
his mother-Ei Jon-A yesterday call-prf  
Aux/past.3sgE.3sgA(ukan)  
His mother called Jon yesterday'
- b. *\*Nor; deitu zuen atzo bere; amak?*
- c. *??Bere amak atzo JON deitu zuen*  
It was Jon that his mother called yesterday'
- (79)a. *Jonek; bere; ama atzo ikusi zuen*  
Jon-E his mother-A yesterday see-prf  
Aux/past.3sgE.3sgA(ukan)  
'Jon saw his mother yesterday'
- b. *Atzo bere; ama JONEK; ikusi zuen*  
Yesterday his mother-A Jon-E see-prf Aux/past.3sgE.3sgA  
'It was Jon that saw his mother yesterday'
- c. *Nork; ikusi zuen atzo bere; ama?*  
who-E see-prf Aux/past.3sgE.3sgA yesterday his mother-A  
'Who saw his mother yesterday?'

In (78a) we see that the possessive can be coindexed with an object noun, but not with an object question word or a focused noun ((78b) and (78c)). When the possessive occurs within an object phrase, these restrictions do not apply (79).

IX.	-TZA	(abstract)		XVIII.	-TE/-TZE	(abstract)		
	<i>bizi</i>	to live	<i>bizitza</i>	'life'	<i>esan</i>	to tell/say	<i>esate</i>	'saying'
	<i>eriotu</i>	to die	<i>eriotza</i>	'death'	<i>ixildu</i>	to be/become silent	<i>ixiltze</i>	'being silent'
	<i>eman</i>	to give	<i>emaitza</i>	'gift'	<i>etorri</i>	to come	<i>etortze</i>	'coming/arrival'
X.	-KUNDE	(concrete)						
	<i>ibili</i>	to walk	<i>ibilkunde</i>	'peregrination'	Suffixes I and III, <i>-tzaile</i> (-er) and <i>-gile</i> (-er), derive agentive nominalizations. These are the only two suffixes in the list above that really qualify as concrete, because they denote humans or animated entities. In these cases, the argument that has been 'absorbed' is the Agent <sup>9</sup> .			
XI.	-KUNE	(abstract)						
	<i>egin</i>	to do/make	<i>egikune</i>	'action'	Suffixes II, IV, V, VI, X, XIII, XXVI, and XVII <i>-ari</i> , <i>-aire/-aira</i> , <i>-kin</i> , <i>-kizun</i> , <i>-kunde</i> , <i>-keta</i> , <i>-era</i> , and <i>-pen</i> , derive nouns that are more or less concrete. Qualifying a derived nominal as concrete or abstract is relative. As mentioned in Chapter 2, it seems that nominal systems are mapped onto a continuum that goes from concrete to abstract. This continuum interacts with other parameters such as event, result, and complex event, creating different readings. How these readings come about is what we try to explore here. Some of the pertinent factors appear to be encoded, either in the lexical entry or in the suffix. Others come about in the syntax. We will see later that derived nouns that are concrete can also be result nominals, and that abstract nominals can be either result or event nominals. The distinction is still relevant, because			
XII.	-KUNTZA	(abstract)						
	<i>ikasi</i>	to learn	<i>ikaskuntza</i>	'study/learning'				
XIII.	-KETA	(concrete)						
	<i>erosi</i>	to buy	<i>erosketa</i>	'purchase'				
XIV.	-KERA	(abstract)						
	<i>sinistu</i>	to believe	<i>siniskera</i>	'belief'				
	<i>ebaki</i>	to cut	<i>ebakera</i>	'cut/way of cutting'				
XV.	-ASUN	(abstract)						
	<i>maite</i>	to love	<i>maitasun(a)</i>	'love'				
XVI.	-ERA	(concrete)						
	<i>irten</i>	to leave/get out	<i>irteera</i>	'exit'				
	<i>esan</i>	to tell/say	<i>esaera</i>	'expression/idiom'				
XVII.	-PEN	(concrete)						
	<i>erosi</i>	to buy	<i>erospen</i>	'purchase'				

*iruditzen zait*

seem-ger Aux/3sgA.1sgD

'Mikel's giving money to his/my mother seems very well to me'

In (81a) we see a simple clause with the three-argument verb *eman* 'give' in which Mikel, the subject, is marked with ergative case, *dirua*, the direct object, with absolute case, and *ama*, the indirect object, marked with the dative case. These three cases are reflected in the form of the auxiliary.

In (81b) we see the nominalized version of (81a), 'Mikel's giving money to the mother.' It does not have the auxiliary form, nor, therefore, the markings on the verb that reflect the arguments. The verb has been nominalized through the suffix *-te* and it has taken the absolute case in order to mark the fact that it has become the subject of the verb in the main clause, *irudi* 'seem'. Note, though, that the nominal arguments of what was a verb in (81a) have kept their case markings: ergative for *Mikel*, absolute for *dirua*, and dative for *ama*.

Also, the nominal built with the suffix *-te* will be able to receive almost any case marking, while the other suffixes have more restrictions. These restrictions may come from semantic and pragmatic factors. But we will be looking at these issues later on.

The next section is devoted to the typology of nominalizations in Basque: their distribution and characteristics.

### 3.2 Typology

As we did earlier for NGPs in English we will discuss the distribution and characteristics of nominalizations in Basque. The section is divided into two subsections one dedicated to the distribution and the other to the characteristics of nominalizations and derived nominals. Special attention is paid in section 3.2.2 to *-te* nominalizations.

#### 3.2.1 Distribution

In this section I examine the distribution that derived nominals and nominalizations in *-te* exhibit. These nominalizations seem to behave as DPs with respect to their distribution. The nominalized form receives the case that corresponds to it, depending on the role it plays in the main clause. Note though, that at the same time the *-te* nominalization is marked for case as if it were a NP, the elements inside its phrasal domain receive the case they would have had, had the nominalized clause be a finite clause. Most of the examples from now on are taken from Artiagoitia (1992), Ortiz de Urbina (1989), and Goenaga (1985).

This structure is particularly interesting because the verb in the main clause allows only two arguments. These are 'father' as the subject and 'me' as the object. The nominalized verb *hitzegeita* 'speaking' is in the absolute case because it is some kind of secondary object of the matrix verb *utzi nau* 'allowed me'. Absolute case is the only possible marking for this nominal.

**(v) Complement to verbs like *lortu* 'succeed in'**

*Amak ez du lortu umeak jatea*  
mother-E neg Aux/3sgA.3sgE succeed child-E eat-nom-A

'Mother hasn't succeeded in the child eating'

*Amak ez du lortu txartela*  
mother-E neg Aux/3sgA.3sgE obtain ticket-A  
'Mother did not obtain the ticket'

The nominalized form *jatea* 'eating' is in the absolute case because it is the object of the matrix verb *lortu du* 'has succeeded'. The subject of the nominalized phrase, *umeak* 'the child,' is marked with ergative case because it is the subject of the nominalized form *jatea* 'eating'.

**(vi) Complement to verbs like *ahaztu* 'forget'**

*Ahaztu zait idaztea*  
forget-prf Aux/3sgA.1sgD write-nom-A  
'I forgot to write'  
*Ahaztu zait liburua*  
forget-prf Aux/3sgA.1sgD book-A  
'I forgot the book'

This is a control structure where the subject of the matrix clause and the subject of the nominalized clause have to be coindexed. *Idaztea* 'writing' is in the absolute because it is the object of the matrix verb *ahaztu zait* 'I forgot'.

**(vii) As adverbial phrases**

*/Gorka gelan sartzean, alde egin*  
Gorka-A room-loc enter-nom-loc away do-prf  
*genuen*  
Aux/3sgA.1plE  
'As soon as Gorka entered the room, we left'  
*Arratsaldean alde egin genuen*  
afternoon-loc away do-prf Aux/3sgA.1plE  
'In the evening, we left'

**(ii) They require case marking**

*Aita hain ondo izateak arritu*  
father-A so well be-nom-E surprise-prf  
*nau*  
Aux/1sgA.3sgE  
'Father's being so well surprised me'  
*\*Aita hain ondo izate arritu nau*  
*Jonek arritu nau*  
Jon-E surprise-prf Aux/1sgA.3sgE  
'Jon surprised me'  
*\*Jon arritu nau*

Nominalizations, like other noun phrases have to be marked for the case that is required by its  $\theta$ -role. They thus enter into an agreement relationship with the main verb.

**(iii) They require agreement in the main clause**

*\*Aita hain ondo izatea arritu*  
father-A so well be-nom-A surprise-prf  
*nau*  
Aux/1sgE.3sgA

Since nominalizations play the same role in regard to the inflected matrix verb as other noun phrases do in regard to the inflected verb in the sentence, they require agreement as any other noun phrase. In the case above, the nominalization is the subject of *arritu nau* 'surprised me' and thus requires ergative case marking. The sentence is ungrammatical because the nominalization is marked for the absolute, instead of the ergative.

**(iv) They can take a genitive modifier.**

*Amaren erosketa ona izan da*  
mother-G buy-nom good be-prf Aux/3sgA  
'Mother's purchase/buying has been good'  
*Amaren janaria ona da*  
mother-G food-A good be/3sgA  
'Mother's food is good'

This genitive can represent the subject, the object or both as we will see later.

**(v) They can have only one object argument in the genitive**

Note that the example above is a zero-derivation example *joan-etorriak* 'goings and comings'. The bare form of the verbs *joan* and *etorri* is taken, and the determiner, absolute case, and plural markers are simply added to them. Other derived nominals can also pluralize, except for nominalizations in *-te*. Nominalizations in *-te* never pluralize.

**(viii) Any DP-internal PP modifier must take -ko.**

*Aitaren gaurko etortzea/etorrera*

Father-G today-G come-nom-A/come-nom-A

'Father's arrival of today's'

*Aitaren gaurko berripapera*

father-G today-G newspaper-A

'Father's newspaper of today's'

This characteristic holds for every member of the nominal system.

**(ix) Word order inside the DP is rather fixed, unlike sentences.**

*Aita dator gaur etxera*

Father-A come/3sgA today house-adl

*Aita gaur dator etxera*

Father-A today come/3sgA house-adl

'Father comes home today'

*Aitaren gaurko etorrera*

Father-G today-G come-nom-A

*\*gaurko aitaren etorrera*

'Todays' father's arrival'

*Aitaren gaurko berripapera*

father-G today-G newspaper-A

'Father's newspaper of today's'

*\*gaurko aitaren berripapera*

**(x) Extraction from the DP is not allowed.**

*Jonek espero du amaren etortzea*

Jon-E await-prf Aux/3sgA.3sgE mother-G come-nom-A

'Jon hopes for/awaits the arrival of his/my mother'

*\*Noren espero du [t etortzea]?*

'Whose arrival does he await?'

*\*Noren espero du [t dirua]?*

'Whose money does he await?'

But the whole DP can be fronted:

*[Noren etortzeal espero du Jonek]?*

'Whose arrival does he await?'

they can take adjectives only if headed by a demonstrative and not the definite determiner.

The only two properties that distinguish *-te* nominalizations from purely verbal constructions are the fact that word order tends to be fixed (but we saw that it is not clear to what degree free word order is a characteristic of finite clauses), and the fact that they do not exhibit certain restrictions that verbal, i.e. tensed constructions exhibit, such as the restriction on 1<sup>st</sup> person dative arguments. I elaborate on these observations later in the chapter.

In the next section we concentrate on *-te* nominalizations in Basque, which were the focus of my fieldwork.

### 3.2.3 Nominalizations in *-te*

Even though nominalizations in *-te* differ little from other nominalizations (by suffix or zero-derivation) in their distribution, they show important differences in their properties.

They require a determiner (i), but if the determiner is a demonstrative their behavior varies. They always require case marking and agreement (ii and iii). They can be modified by a relative (vi), but they cannot pluralize (vii), even when they seem to be totally “nominal”. If they take a DP-internal modifier, the modifier has to be marked by *-ko* (viii), but

they do not participate in compound formation (xi). They can take adjectives when the demonstrative is used (xii), but take adverbs if the determiner is the definite article.

(82) *Aitonak isilik lan egitea harigarri egiten zitzaidan*  
Grandfather-E silently work do-nom-A surprising do-ger Aux/past.3sgA.1sgD  
'Grandfather's silently working was surprising to me'

Note that *aitonak* 'grandfather' is in the ergative case. It is marked as subject of the nominalization *lan egitea* 'working'. The adverb would not be possible if the subject had been in the genitive.

If they take genitive modifiers, they will follow the same restrictions on genitives as do the other derived nouns (iv and v). But the interaction between genitive modifiers or complements and other complements will vary.

Like other result nominals, they do not allow the interaction of the genitive with certain complements.

(83)a. *\*Zure denda hartan liburu erotea*  
your store-A that-loc book-bare buy-nom-A

(85)a. *Aitaren berebilaren erosketa oso ona*  
 father-G car-G purchase-A very good  
*izan zen*  
 be-prf Aux/3sgA.past  
 'Father's purchase of the car has been good' ('Father's car's purchase..')

b. ??*Aitaren berebilaren erostea oso ona izan zen*

For many speakers, (85b) is ungrammatical. The only difference between (85a) and (85b) is the type of nominalization used: a derived nominal in *-keta* in (85a) (*erosketa*); and a derived nominal in *-te* (*erostea*) in (85b).

Nominalizations in *-te* do not allow extraction from the DP (x), but they do allow extraction when the arguments are case-marked, i.e., when the structure is more "verbal".

(86)a. *Jonek espero du amaren etortzea*  
 Jon-E await-prf Aux/3sgA.3sgE mother-G come-nom-A  
 'Jon hopes for/awaits the arrival of his/my mother'  
*\*Noren espero du [t etortzea]?*  
 'Whose arrival does he await?'

b. *Jonek espero du amak Peiori*  
 Jon-E await-prf Aux/3sgA.3sgE mother-E Peio-D  
*dirua ematea*  
 money-A give-nom-A  
 'Jon hopes for/awaits mother's giving money to Peio'  
*Nori espero du Jonek [amak t dirua ematea]?*  
 'To whom does Jon hope mother's giving money?'  
*Zer espero du Jonek [amak Peiori t ematea]?*  
 'What does Jon hope mother will give to Peio?'

The contrast between (86a) and (86b) shows that extraction is allowed from a *-te* nominalization when the arguments are case marked.

As mentioned above, the direct object of the nominalization, if the nominalization is a DP, may appear in the indefinite form, without case marking.

(87)a. *Bizarra moz-tea erabaki dut*  
 beard-A cut-nom-A decide-prf Aux/3sgA.1sgE  
 b. *\*Bizar moztea erabaki dut*  
 beard-bare cut-nom-A decide-prf Aux/3sgA.1sgE  
 'I have decided to cut my beard'  
 c. *\*Lana erabaki dut*  
 work-A decide-prf Aux/3sgA.1sgE  
 'I decided work/job'

c. *Etxean gelditu behar du-gu*  
 house-def-in stay-prf must Aux/3sgA.1pIE  
 'We must stay home'

In (88a) we see the nominalized form in *-te* of the verb *egin* 'do' or 'make'. This nominalized verb has a direct object, *lan hori* 'that work/job,' marked in the absolute case and a subject, *guk* 'we,' marked in the ergative case.

In (88b) the verb has taken the form of the participial adjective and is inflected in the absolute plural as is its 'object', *eskutitzak* 'letters'. This form of the verb could not have a subject that is different from the subject of the main verb.

In (88c), (88d) and (88e) we see structures where the subject is controlled and the verb is in its root or perfective/infinitival form.

From all the types of nominalizations listed in Section 3.1.3, derivations with suffixes other than *-te*, zero-derivations, and a subset of *-te* nominalizations occur at the lexical level. Only a subset of the nominalizations in *-te* may occur at the syntactic level.

In Section 3.3.1.1, we will be looking at an analysis that rejects this difference in nominalizations in *-te* and splits them on their phrasal status. Nominalizations in *-te* might be considered CPs or DPs. Both

types of phrases have an inflectional head. These inflectional heads are simply of a different type in each case. This partially solves the endocentricity issue mentioned in Chapter 2, but does not account for some of the differences between the DP-like nominalizations in *-te* and their counterparts derived through other suffixes or zero-derivation, mentioned in Section 3.2.2.1.

We also present a Double Insertion Hypothesis approach, by which the derivation of *-te* nominalizations may occur either at the D-structure level or in S-structure. This account has similar limitations to those mentioned above. In addition, it does not solve the endocentricity problem.

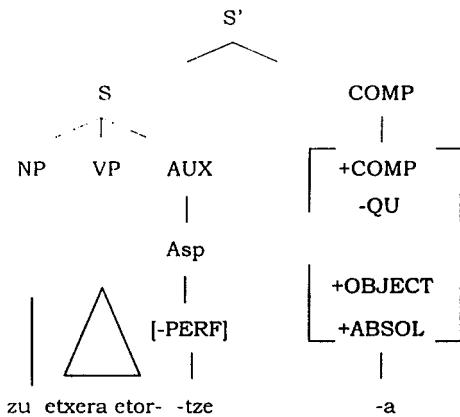
Section 3.3.2 will be devoted to control phenomena. In section 3.3.3 I present Grimshaw's (1990) argument structure hypothesis and apply it to the Basque data. Her account has many advantages in regard to DP-like nominalizations, but leaves totally unexplained the change of category of CP-like nominalizations.

In section 3.3.5, I will comment on the predictability issues that this type of data suggests.

tense interpretation, but it cannot be situated in the past with respect to the main clause, as can be seen in (91).

(91) \*Andonik idi bat hegan ikustea esan  
 Andoni-E bull-bare one flying see-nom-A say-prf  
*digu*  
 Aux/3sgA.1plD.3sgE  
 "Andoni told us seeing a steer flying"

Goenaga proposes the following structure for nominalized clauses:



He concludes that *-te* is the inflectional head, the auxiliary.

Concentrating on giving features to the matrix verb, Goenaga argues that the distribution of the subordinate in *-te* can be accounted for by the features [-PROPOSITIONAL] and maybe [+ACTION] in the matrix verb.

Goenaga also notes that this type of nominalizations can behave like regular NPs. In this context, he considers *-te* to be a nominalizing suffix like any other (-KETA, -ERA, etc.).

Zabala (1995) assumes Abney's (1987) D-IP hypothesis in Basque. She notes that the phrase whose head is suffixed with *-te* shows ambiguity in its aspectual reading:

(92) Miren berandu heltzea harrigarria da  
 Miren/A late arrive-nom-A surprising be/3sgA  
 "The fact that Miren arrives/has arrived late is surprising"

Both the perfective and imperfective reading are obtained.

Nominalizations can also express modality:

(93) Gorrotu dut Miren ixildu  
 hate-prf Aux/3sgA.1sgE Miren/A get quiet-prf  
 ezin izatea  
 impossible be-nom-A  
 "I hate the fact that Miren cannot be quiet"

She argues that the only difference between finite and non-finite clauses in Basque is that in the latter, the features for T and Agr are weak and do not get realized until LF.

Ortiz de Urbina (1989) follows a similar line of thought and argues that only tenseless embedded clauses that are in a position where case is assigned can assign case to their subject. He posits that *-te* is responsible for derived nominals created in the lexicon. He argues that, in this case, the suffix is derivational. *-te* is also a tenseless inflection that is attached to V. This position leaves the contrast between (94b) and (94c) unexplained.

Ortiz de Urbina notes that the fact that case can be assigned to the subject of a tenseless clause only when the clause is itself in a position to which case has been assigned extends to participial complements. I take the examples from Ortiz de Urbina.

(95)a. */aitak/[e] semea lepoan eramanak] poztu*  
 father-E son-a shoulder-on carry-E gladden-prf  
*nau*  
 Aux/3sgA.3sgE

'I'm glad at the father's/[e] having carried the son on his shoulder'

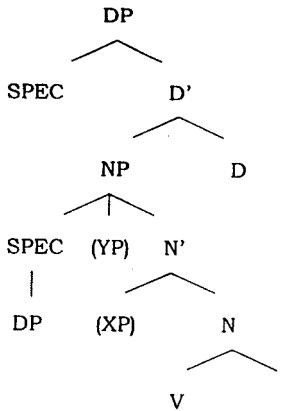
b. */liburu hau- /Axularrek aspaldian idatzia] zen*  
 book this-A Axular-E long ago write-A was  
 'This book was written by Axular long ago'

c. *bankua [polizia konturatut] gabe lapurtu*  
 bank-A police-A notice without steal-prf  
*zuten*  
 Aux/past.3plE.3sgA  
 'They robbed the bank without the police noticing'

(96)a. *guk ez dakigu [\*bera/\*jon/[e] nora joan]*  
 we-E neg know-1pl he-A/Jon-A where-adl go-inf  
 'We do not know where (\*he/\*Jon) to go'

b. *Peruk [\*bera/\*Jon/[e] etxera joan] nahi*  
 Peru-E he-A/Jon-A home-adl go-inf want  
*du*  
 Aux/3sgE.3sgA  
 'Peru wants (\*he/\*Jon) to go home'

Ortiz de Urbina, following Raposo (1987), argues that the contrast between (95) and (96) shows that only tenseless clauses that have been assigned case can assign case. But these data do not support this position. For one thing, (96a) is a control structure and, thus, for independent reasons the subordinate clause cannot have a lexical subject. Example (96b) shows that a lexical case-marked subject is impossible for infinitival clauses, but not for all tenseless clauses. In this environment a nominalized clause with *-te* can have a lexical case-marked subject. Let me repeat (94c) here:



Artiagoitia (1992) applies five tests to distinguish DPs from sentences in Basque. Examples are taken from Artiagoitia.

**(i) DPs may bear ergative case, Ss cannot be subjects of transitive verbs.**

(97)a. Ainhoak izugarri kezkatzen nau  
 Ainhoa-E terribly worry-ger Aux/1sgA.3sgE  
 'Ainhoa worries me terribly'

b. \*Datorren urtean zer eginak kezkatzen  
 next year-loc what do-E worry-ger  
 nau  
 Aux/1sgA.3sgE  
 'What to do next year worries me'

c. Ainhoa etortzeak kezkatzen nau  
 Ainhoa-A come-nom-E worry-ger Aux/1sgA.3sgE  
 'Ainhoa coming worries me'

Item (97c) shows a nominalization as subject of the matrix clause.

According to Artiagoitia, this is impossible for an infinitival clause (97b), though recall (95a), which shows the same environment and is considered grammatical by Ortiz de Urbina (1989).

**(ii) Sentential subjects are possible with unaccusatives and copulative verbs, but they are incompatible with wh-movement of a complement unless extraposed.**

(98)a. Ainhoak /garagardoa ekar  
 Ainhoa-E beer-A bring-bare  
 dezala] beharrezko da  
 Aux/subj.3sgA.3sgE-comp necessary is  
 nire ustez  
 my belief-instr  
 'That Ainhoa should bring beer is necessary in my opinion'

b. \*Noren ustez da /Ainhoak garagardoa ekar  
 who-G belief-instr is Ainhoa-E beer-A bring  
 dezala] beharrezko?  
 Aux/subj.esgA.3sgE-comp necessary

c. *Lehendakariak [presaen askapena] eta*  
 President-E prisoner/G.pl liberation and  
 [gobernuak suetena negoziaztea]  
 government-E cease-fire-A negotiate-nom-A  
 eskatu du/ditu  
 ask-prf Aux/3sgE.3sgA (3plA)

'The president demanded the liberation of the prisoners  
 and the government's negotiating a cease-fire'

Example (101a) shows that a tensed clause does not coordinate with a DP. Example (101b) shows that a nominalization with *-te* does not coordinate with a tensed clause. But in (101c), we see that a nominalization in *-te* does, indeed, coordinate with a DP.

**(iv) Coordination of sentences does not trigger plural, but coordination of DPs does or, at least, may.**

(102) a. *Gobernuak eta ETAk [presaen askapena]*  
 government-E and ETA-E prisoner/G.pl liberation  
*eta [indarkeriaren amaiera] negoziatuko*  
 and violence-G end-A negotiate-fut  
*dute/dituzte*  
 Aux/3plE.3sgA(3plA)

'The government and ETA will negotiate the liberation of

the prisoners and the end of violence.'

b. *Gobernuak eta ETAk [suetena noiz hasi]*  
 Government-E and ETA-E cease-fire-A when start-prf  
*eta [presaok noiz askatu] negoziatuko*  
 and prisoner/A.pl when liberate-prf negotiate-fut  
*dute/ \*dituzte*  
 Aux/3plE.3sgA (\*3plA)

'The government and ETA will negotiate when to start the  
 cease-fire when to free the prisoners'

c. *Gobernuak eta ETAk [presaok askatzea]*  
 government-E and ETA-E prisoner/A.pl liberate-nom-A  
*eta harmak betirako uztea negoziatuko*  
 And weapon/A.pl for ever quit-nom-A negotiate-fut  
*dute/dituzte*  
 Aux/3plE.3sgA (3plA)

'The government and ETA will negotiate freeing the  
 prisoners and putting down the weapons'

In (102b), where two clauses are being coordinated, the plural form of the auxiliary (*dituzte*) is not possible. The plural is possible in (102a) and (102c), in which two nominalizations are coordinated.

### 3.3.2 Control Phenomena

Zabala (1995) argues that the empty categories present in *-te* nominalizations in Basque do not belong to the category PRO, but pro.

The same point was made by Rebuschi (1989):

*It thus appears that control and lack of government are two distinct issues: since full Case-marked subjects are possible, the unrealized subjects...must be treated like (small) pro's not PROs, although their semantic interpretation is identical to that of PRO in English and many other languages* [p.19].

(104) a. /Zuk Peruri dirua ematea/ nahi  
 You-E Peru-D money-A give-nom-A want-prf  
 dut  
 Aux/1sgE.3sgA  
 'I want you to give money to Peter'

b. /e<sub>i</sub> e<sub>k</sub> e<sub>h</sub> ematea/ nahi dut  
 give-nom-A want-prf Aux/1sgE.3sgA  
 'I want you to give it to him/her'

c. Itziarrek /mendira joatea/ erabaki  
 Itziar-E mountain-adl go-nom-A decide-prf  
 du  
 Aux/3sgE.3sgA  
 'Itziar decided to go to the mountain'

d. Itziarrek /gu denok mendira joatea/ erabaki  
 Itziar-E we-A all mountain-adl go-nom-A decide-prf  
 du  
 Aux/3sgE.3sgA  
 'Itziar decided that we all go to the mountain'

The examples in (104) show that the empty category has to be pro and not PRO. But:

(105) Nahi duzu /lagunzea/?  
 want-prf Aux/2sgE.3sgA help-nom-A  
 'Do you want (me) to help you?' (only possible reading)

Example (105) shows that the subject of the nominalized clause is controlled. But as noted by Joseph (personal communication), this utterance may not be exhibiting syntactic control, rather it may be explained through pragmatic factors.

These empty categories do not have to be controlled, but can be, and do not need to have arbitrary reference. They show obviation phenomena (typical of pro), and the arguments can be phonetically realized. Also the empty category can correspond to subjects, objects, and indirect objects.

Grimshaw's basic assumptions (which we already briefly mentioned in Chapter 2) are:

- A-structure is a structured representation which represents prominence relations among arguments.

*announce(Agent(Goal(Theme)))*

In the example above Agent is the most prominent role and Goal is more prominent than Theme.

- Some thematic information is encoded in a-structure.
- The concept of an external argument can be explained by prominence relations in a-structure.
- There is a distinction between grammatical arguments and semantic participants. Each entry has a lexico-semantic representation (lcs, lexical conceptual structure). The lcs of a verb encodes, among other things, the participants in the activities or states described by the verb. Not all entries have the ability to project their lcs into a-structure.
- The a-structure of lexical items varies across syntactic categories.

For Grimshaw (1990), some derived nouns have an a-structure and others do not, even though they all have a lcs inherited from the verbs from which they come. Only nouns denoting complex event

structures will have an a-structure. Result and simple event nominals do not. Nouns are defective theta markers in English, and this is why even complex event nominals need prepositions to mark their arguments.

She distinguishes the two readings of a noun such as 'assignment' in English, using as a test tool the adjective 'constant' which forces the complex event reading (English examples taken from Grimshaw):

(107) a. *The assignment is to be avoided.*  
b. \**The constant assignment is to be avoided.*  
c. *The constant assignment of unsolvable problems is to be avoided.*  
d. *We constantly assign \*(unsolvable problems).*

The indefinite determiner, the numeral 'one,' and demonstratives are compatible only with result nominals. The definite determiner 'the' occurs with both.

(108) a. *They studied the/an/one/that assignment*  
b. *They observed the/\*an/\*one/\*that assignment of the problem.*

Complex event nominals do not pluralize, while result nominals do.

(109) a. *The shooting of rabbits is illegal.*  
b. \**The shootings of rabbits are illegal.*

This distinction between complex event and simple event nominals in Basque is difficult to capture because lexical entries of the 'trip' type are rare and are often loans. Most of the event-denoting nouns are derived (*bidaia* 'trip,' *lasterketa* 'race,' *azterketa* 'exam,' *gertaera* 'event') and are ambiguous between a result and an event reading.

We will assume, then, that the relevant distinction in Basque is between complex event and result nominals. Grimshaw also predicts that gerundive nominals are going to be complex event nominals. This prediction is borne out in Basque by the *-te* nominals. They do not pluralize, they take adjectives of the 'constant' type, and they exhibit event control.

(114) a. *Aitaren istorioen kontatzea [giroa*  
 father-G story/G.pl tell-nom-A ambiance-A  
*obetze]*  
 improve-nom-prol  
 'Father's telling of stories to make things better'  
 b. *Aitak istorioak kontatzea [giroa*  
 father-E story/A.pl tell-nom-A ambiance-A  
*obetze]*  
 improve-nom-prol  
 'Father's telling stories to make things better'

Nominalizations and Derived Nominals in Basque can be mapped into three continua:

**Module**

Syntactic  $\longleftrightarrow$  Lexical

**Category**

CP/IP  $\longleftrightarrow$  DP

**Argument Structure**

Event  $\longleftrightarrow$  Result

To summarize, certain zero-derived nominals (*bildur*, 'fear'), nominals derived through certain suffixes (*kondaketa*, 'telling'), and *-te* nominalizations can be both result and event nominalizations.

What sets *-te* nominalizations apart is the fact that their external behavior in the syntax varies depending on which reading they obtain.

When *-te* nominalizations obtain a result reading they behave like any other NP. They are defective theta markers and present all the properties and the distribution listed above as DP-like.

When *-te* nominalizations obtain the event reading, they behave almost (and here is where the differences with English can be seen) like VPs or CPs. They assign case like any other verb. They do not fulfill some of the properties of DPs (do not participate in compound formation,

d. the process of V

Nominal *-te* nominalizations may obtain interpretations a., b. and

c. Verbal *-te* nominalizations may obtain interpretations a., b. and d.

Nominal *-te* nominalizations can appear with a demonstrative or with a determiner. Nominalizations with a determiner can obtain only interpretations a. and c.

The examples cited in (94) show the contrast between nominal *-te* nominalizations headed by a determiner and those headed by a demonstrative. The environments in which these variants are not possible are those where we cannot have a 'fact of' or 'way of' interpretation. I repeat the examples here for convenience:

(94)a. *Gorroto dut*      *Kepak jakin*      *gabe*

hate   Aux/3sgA.1sgE Kepa-E know-bare without

*hitz egite*      *hori*

speak-nom-bare that-A

'I hate that Kepa's speaking without knowing'

b. \**Kepak lasai hitz egite*      *hori nahi*

Kepa-E calm speak-nom-bare that-A want-prf

*dut*

Aux/3sgA.1sgE

'I want that calm speaking of Kepa's'

c. *Kepak lasai hitz egitea*      *nahi dut*

Kepa-E calm speak-nom-A want Aux/3sgA.1sgE

'I want Kepa to speak calmly'

d. \**[Kepak gelara sartze horretan] guztiak*

Kepa-E room-adl enter-nom-bare that-loc all-A.pl

*ixildu*      *ziren*

become silent-prf Aux/past.3plA

\*'At that entering of Peter's, they all became silent'

e. \**Kepak ahaztu du*      *[e]dirua ekartze*

Kepa-E forget-prf Aux/3sgE.3sgA [ money-A bring-nom

*hori]*

that-A]

\*'Kepa forgot that bringing money (of himself)'

The restriction is semantic and not syntactic. Verbs, depending on their lcs and their interpretation will be able to fit in certain syntactic environments and others will not. However, syntactic restrictions are also at play. For example, (94e) shows a syntactic environment where control is required. Let's assume that a main clause whose verb is of the type *ahaztu* 'forget' subcategorizes by default for a clause. If this is so, then by default the nominalization that will be selected will be verbal. But, there are also pragmatic factors that come into play. What could it

mean to say something like 'I forgot that speaking of John's?': Perhaps only 'I forgot his style of speaking', not 'I forgot that he spoke'.

The difference of distribution between nominal *-te* nominalizations and verbal *-te* nominalizations is dependent upon both syntactic and semantic factors. Verbal *-te* nominalizations will be required in the context of a complement to a verb that requires a clause as a complement. In this case, these nominalizations will be in complementary distribution with other tenseless and tensed subordinate clauses.

When the nominalization is occupying a position that could be occupied by a NP, the distinction is semantic. We will prefer a nominal *-te* nominalization when the required interpretation is more factive or manner-like, and a verbal *-te* nominalization in the other cases.

The type of restrictions that apply across the board and in all contexts on the choice of a more verbal or more nominal structure concerns tense or aspect. Recall example (90), which I repeat here for convenience:

(90) a. *Andonik idi bat hegan ikusi*  
Andoni-E steer-bare one flying see-prf  
*duela esan digu*  
Aux/3sgA.3sgE-comp say-prf Aux/3sgA.1plD.3sgE  
'Andoni told us that he saw a steer flying'

In a context like that shown in (90) it is impossible to have a nominalization:

(115) \**Andonik idi bat hegan ikustea*  
Andoni-E steer-bare one flying see-nom-A  
*esan digu*  
say/prf Aux/3sgA.1plD.3sgE  
'\*Andoni told us that seeing a steer flying'

A nominalization is impossible here because the tense interpretation of the nominalization does not fit the needs of the main clause.

Once again, Grimshaw's (1990) theory allows us to make the correct distinction here. We will be able to predict the occurrence of a *-te* nominalization by combining the criteria presented in 3.3.3 and other semantic criteria. These semantic criteria affect the choice of one nominalization or the other exactly because the Ics of the base verb is

If we follow Grimshaw (1990) we would probably qualify these nominalizations as simply verbal, as having a complex event structure because they are verbal, like the verbal *-ing* in English. Even though this is correct, as we saw above, these constructions can not be considered simply verbal.

These structures are verbal, but they have lost their inflectional head, the auxiliary. The arguments are case-marked and that morphological marking tells us that they have projected an argument structure from the corresponding verb's ICS and that, unlike derived nominals, they are not defective theta markers. The fact that they are not defective theta markers does not mean that they are verbs. As was said above, they seem to be almost verbs but not quite.

As opposed to NPs, verbal nominalizations exhibit similar freedom in word order as do tensed clauses. As we said above, the arguments are case-marked. As has been mentioned already, though, they do not seem to have the same opportunities to define tense and aspect as tensed, or even infinitival clauses have. Another interesting contrast with tensed clauses that nominalizations exhibit is the one noted by Laka (1993), and to which we have referred before.

In tensed clauses, when a verb has three arguments, the absolute argument has to be third person. This restriction does not apply in nominalizations:

(118) a. \**Zuk ni etsaiari saldu naiozu*  
you-E I-A enemy-D sell-prf Aux/1sgE.1sgA.3sgD  
'You sold me to the enemy'  
b. *Gaizki iruditzen zait* *[zuk ni etsaiari saltzea]*  
wrong seem-ger Aux/3sgA.1sgD you-E I-A ennemy-D  
sell-nom-A  
'Your selling me to the enemy seems wrong to me'

This kind of contrast regarding interactions between the person of arguments is also manifested in the Romance languages:

(119) a. *Te lo di*  
you/D it/Acc give/pret  
'I gave it to you'  
b. \**Te le di*  
you/Acc s/he/D give/pret  
'I gave you to him/her'

Whichever interpretation/translation of (119b) from Spanish one uses, the interaction of these two clitics is impossible. Artiagoitia (1997),

difficulties of the task vary depending on which two languages are used. This means that, if the two languages are typologically close, one will get more 'free rides' i.e., the cases where the target language is generated correctly by accident, either because of typological similarities or because of default conditions. The number of 'free rides' will also depend on the structure that the linguist is trying to account for.

When the structure that needs to be analysed is a -ing and the source language is English, for example, great difficulties arise at the parsing level.

The Logos Machine Translation system uses a home-grown language called SAL. This language not only encodes semantic and syntactic features of a lexical entry, but it is also the language used for writing rules and encoding commands to various programs written in C++ and Java.

Thus, words are replaced by codes. These are some examples (BOS stands for "beginning of sentence"):

BOS	20	1	1
Other	4	81	23
forms	1	40	2
of	11	52	2
storage media	1	77	2

As can be seen, the entry 'forms' gives three options to the parser: a plural noun; a transitive verb; and an intransitive verb.

Features are encoded at the lexical level in two ways. One type of encoding is shown above: the entry is practically "translated" into a numeric language. Other features are hidden and stored in "scons". These are containers attached to lexical entries. Each entry has 100 scons. Some of these scons are filled at the lexical level, while others are used as parsing and generation progresses. In the course of the derivation, scons may contain both source and target information.

The image below shows an example of the data structure:

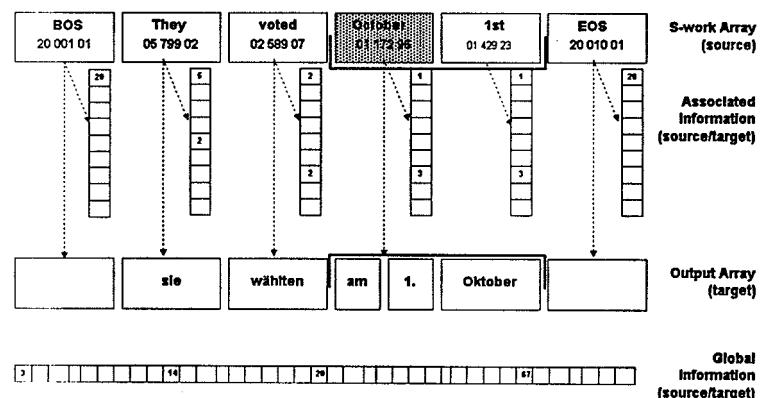


Figure 1. Linguistic Data Structures in Logos

account for phenomena of categorial change, such as the one involved in nominalizations, this observation is not sufficient.

When an *-ing* is to be parsed in English, the most difficult distinction to make is that among an adjective, a verb (transitive and intransitive), and a noun. Some *-ings* have been lexicalized in English, like *bearing* or *clipping*. But even these have Process Noun counterparts, which, being mixed categories as they are, exhibit verbal behavior. Thus, lexicalization helps to a certain extent, but does not solve the problem. The distinction between adjective and verb is not obvious either:

(121)a. *Those guys are annoying fellows*  
b. *Those guys are annoying neighbors*

In regard to the distinction between verbs and nouns, the parser will by default tend to analyse *-ings* as verbs. As parsing and generation proceeds, the system will try to analyse the structure further in order to make a decision and to keep informing the target language of the clues collected, in case the target language wants to override some of the decisions.

As a consequence of the analysis of Basque nominalizations presented in the sections above, a computational solution can be provided for such structures.

In the following section I present a sketch of a computational implementation of Basque nominalizations into a Machine Translation system.

### 3.5.1 Parsing Basque Nominalizations

Lexicalized derived nominals will be encoded at the lexical level. Leaving aside the suffix *-te*, there are at least 18 suffixes that derive nominalizations in Basque. In addition, we need to account for the zero-derived nominals. It wouldn't be appropriate to store all these forms in the dictionary. We thus need to have a strong morphological component in order to "un-derive" these forms and characterize them properly.

The pieces of information that will have to be encoded in the feature database and in the rule database in order to parse the structures correctly will be:

- The verb from which they derive.
- The lcs of the verb from which they come.
- The number and type of arguments of the verb from which they derive.

Let's suppose that we assign value 1 for unaccusatives, value 2 for transitives, 2b for unaccusatives plus dative, and 3 for ditransitives. By

same rules and features can be used for derived nominals and for nominalizations in *-te*. This extends also to zero-derivation nominalizations.

### 3.5.2 Generation of Basque Nominalizations

Generation turns out to be more difficult than parsing. As opposed to parsers that are built for testing theoretical linguistics concerns, parsers in a commercial Machine Translation system need only to encode information and analyse it to the point that is necessary to understand and code the phenomena to correctly generate a target language.

Verbs in Basque will have to be marked for their requirements concerning clausal subcategorization. *Nahi*, 'want,' for example, will have to be marked as pre-verbal and pre-clausal.

When the verb in question is acting as a pre-verbal, we generate an infinitive in Basque for the subordinate verb. When it is acting as pre-clausal, we have the choice between a finite and a non-finite clause. Verbs also have to be marked for this feature. Tense and aspect considerations allow us to make the correct choice.

If we had encountered a verb of the 'want' or the 'be able' type followed by a subordinate clause or an infinitival, we would generate a *-te* nominalization.

When the feature calls for a non-finite clause, depending on other semantic requirements, either a *-te* nominalization, an infinitive, or a bare form will be generated. In all these cases, the phrase-internal arguments will be marked for case.

When we encounter an infinitival in subject position, a *-te* result nominalization, with the arguments marked for case will be generated.

Both in DP and sentential environments, the default setting for the *-te* nominalization will be to have the arguments marked for case. It will, then, by default, be more verbal than nominal.

When the translation demands the arguments to not be case-marked (by the presence of a demonstrative, for example), derived nominals other than *-te* will be preferred.

Some of the issues must be handled in the semantic component of the system. For example, once we have identified an *-ing* in English as being either a concrete noun or a process noun in a limited semantic context, we can generate the correct type of nominalization in Basque right in the semantic module.

Linguistics in general and, more specifically syntactic studies, have concentrated, thus, on the understanding of linguistic knowledge and on language as a system that exists independently from external factors.

Parallel to this position, sociolinguists, anthropological linguists, and ethnographers, among others, insist that there exists a great deal for Linguistics to learn from investigating language in use. Research in language attrition fits into this group of inquiries.

This chapter will serve as the context within which to test my original hypotheses. I adopted as a working hypothesis that Basque nominalizations are a complex structure and that they are thus vulnerable in language contact situations in which simplified structures might be expected to prevail. This hypothesis is tested against the data in Chapter 5.

It is worth noting that other authors have observed attrition phenomena in regard to nominalizations or gerunds in language loss settings. However, if Thomason and Kaufman (1991) are correct, there are no *linguistic* constraints on what can be affected in a language due to contact with another language (though social factors can affect the outcome of contact). Still the fact that nominalizations are documented as being affected in other contact situations means that no one could claim that nominalizations are impervious to the effects of contact.

Denison (1977) mentions the case of a village in North Eastern Italy, Sauris, where the speakers of German, under massive influence from the Romance languages (Italian, Friulian) are using Romance forms in their nominalizations. For example, among German speakers the form [ɔndars] is normal for the notion 'different', but for 'difference' they exhibit only [də dife'rentsa].

Morales (1981) observed an increase in the use of the gerundive forms instead of the infinitive in Puerto Rico. Constructions such as '*la mejor manera de evitar el vicio de fumar es mascando chicle*,' instead of '*mascar chicle*', as found in other dialects such as Peninsular Spanish, ('the best way to avoid the vice of smoking is chewing gum') reached a 42% acceptance rate among the speakers he tested. López Morales (1993) observes another environment where the gerund is replacing the infinitive: '*Nadando es bueno para la salud*,' instead of '*Nadar*' ('Swimming is good for your health'), or '*Lo que hace es comparando muestras*,' instead of '*comparar*' ('What s/he is doing is to comparing samples'). In general, it appears that some varieties of Spanish in the Americas exhibit phenomena of loss of subordinate constructions possibly under the influence of English, at least in the case of Puerto Rican Spanish.

The Recent Filler and Salience strategy specifies that the processor tends to assign the most recent potential filler to an encountered gap.

But unambiguous sentences also show differences in complexity, at least it is so perceived by the speakers. There have been proposals that have analyzed these differences in terms of depth (Ingve, 1960), in terms of amounts of superstructure to be assigned to a word (the nonterminal-to-terminal node metric), etc.

It seems that none of these explanations have yielded clear results and the reader is directed to the literature on complexity and garden path analyses. Complexity exists in language processing (psycholinguistics has demonstrated it), and I will take as a working hypothesis that complexity is a relevant factor in a process of language decay.

For my purposes, and in the context of language attrition, I will assume that complexity comes from ambiguity (more than one possibility of understanding a sentence), from word order variations, and from specificity of marking.

Ambiguity is often due to a situation of polysemy for a word and also to multiple possible subcategorizations and syntactic behaviors. The choice of the right option is often connected also with pragmatic or

conversational factors. The choice between an infinitival clause with or without a determiner in Spanish and the choice between one or another type of nominalization in Basque, for example, are not determined exclusively by syntactic and semantic factors. In order to select the correct option, the speaker needs to know not only the grammar of the language, but also the appropriate environments, contexts and registers to which a grammatical restriction is connected.

Word order differences reveal different interpretations of a sentence. This applies not only to languages with a “fixed” word order like English, but also, as we saw in Chapter 3, to languages that supposedly have a “free” word order, such as Basque. The speaker needs to know not only the available word order options but also their appropriate use.

What I mean by “specificity of marking” is the fact that certain linguistic structures are more “restricted” than others. The speaker has the choice among quite a large number of possible postpositions in Basque, of deverbalization possibilities, and of various subordinated clause constructions. The more “specific” the option, the higher the risk of incurring ungrammaticality. It is known that second language learners tend not to risk committing errors and avoid these typical

#### 4.2 Language and Ethnicity

Against this backdrop of formal approaches to syntax, it is important to keep social functions of language in mind. Language is not only a means of communication, it is also an identity, a past, as Fishman (1977) states:

*In view of the foregoing, it becomes clearer why language is more likely than most symbols of ethnicity to become the symbol of ethnicity. Language is the recorder of paternity, the expresser of patrimony and the carrier of phenomenology (p.25).*

Language is probably the most important factor defining ethnicity. Through ethnic collectivities, in turn, individuals feel augmented beyond their loneliness and their inherent "frailty".

The last forty years have witnessed a rebirth of ethnicity that could not have been foreseen. Ethnicity or, its frequent consequence, 'nationalism,' has been traditionally seen as negative, emotionally-based, and irrational, and often as a dangerous concept. But many minorities embraced ethnic consciousness, claiming their right to national status, a right allotted to majority, established nations but not to minority, non-established nations.

Language has always been one of the banners claimed and used in nationalistic movements. In the case of the Basque community, the language is very important. The Basque word itself that means 'a Basque person' translates literally into 'the one that has the Basque' *euskalduna* → *euskal-dun-a*. There are other signs of identity for being Basque, but this lexical entry has been used by politicians and nationalistic movements as a symbolic argument, among others, for the defense of the language. For some reason, it works.

In most countries in the world more than one language is spoken. This is not surprising. What is surprising, however, is that this reality has not been accepted in most parts of the world. Usually, it is denied, suppressed, or ignored.

Languages have different status. This situation is partly due to policies, partly due to factors intrinsic to the communities themselves. Some languages have fewer speakers than others. Some languages are not official, and they do not have a legal status. Some languages are not written. Some countries, and, thus, their official languages, are more powerful than others. Some languages have the status of international languages; while others do not. The social and political status of a language is, thus, determined by several factors. These factors are

#### 4.3 Language Contact and Language Change

Languages evolve constantly. Even though much is known about linguistic change, questions remain over the direction of those changes, their outcome, as well as what the causes of change are. Labov (1972, p. 160) asks the following questions regarding these phenomena:

1. *Is there an overall direction of linguistic evolution?*
2. *What are the universal constraints upon linguistic change?*
3. *What are the causes of the continual origination of new linguistic changes?*
4. *By what mechanism do changes proceed?*
5. *Is there an adaptive function to linguistic evolution?*

A distinction needs to be made, though, between monolingual and multilingual settings. There is a tendency both to forget certain facts of multilingual phenomena, when analyzing linguistic changes, as well as a tendency to stereotype the problems and issues of the multilingual community.

An example of this is the generalized attitude of tolerance toward change and against purism. Linguists are supposed to observe what occurs in languages without trying to intervene in any way. Even though it is correct that value judgments about languages are unscientific, it is also true that when a language is in danger of extinction certain

measures, which will certainly appear to be attempts at manipulation of the 'natural' evolution of a language, ought to be taken and linguists ought to be involved. What I mean by this is that linguists sometimes have to take a stand and use what they know about language in order to counter the potential death of a linguistic variety<sup>12</sup>.

This is not very different from the tendency of purists who try to keep a language 'clean' and without interference. On the one hand, linguists have a tendency to disagree with this type of position. On the other, they seem engaged in saving languages like Basque. This is a very ambiguous situation. What I am trying to say is that when one observes and studies language change in a monolingual setting (historical change) and when one does it in a multilingual setting, the goals may not be the same.

Thus, my attitude is very clear here in that knowledge serves a purpose and the observation of maintenance or death of a language may give us ways of diagnosing decay as opposed to evolution. This information is relevant to linguistic inquiry in general, but it is also relevant in assisting communities in saving their language. This information should be useful and applied to the measures to be taken in order to protect a language, speakers willing, naturally. Language change can be manipulated up to a certain point. Discovering the

of this concept as a valid linguistic concept has not been established, but the term has spread into the anthropological linguistics and sociolinguistic literature. What is more worrisome is that terms and notions similar to semilingualism appear in official educational documents with negative connotations. Eldeksy et al. (1983) report that the term 'limited English dominant' was used to describe children's language skills in educational legislation in California.

Intuitively, though, this is a useful term to define a certain type of speaker. While more research needs to go into specifying what the term covers, for the present purpose, we will take it to refer to speakers who have difficulties expressing themselves and understanding others with an ease comparable to that of a full speaker. These are speakers who may have been native speakers and have lost command of their language. The term also covers, though, imperfect speakers who have learned the language later in life. The important difference between this type of speaker and speakers of a foreign language is that 'semi-speakers' usually control conversational constraints and rules. These individuals often behave as if they were ordinary members of the linguistic community. Their deficiencies frequently go unnoticed because of their knowledge of the interactional norms of the linguistic community. Identifying these semi-speakers is a challenge for the fieldworker, since s/he needs to know how reliable the data are.

Some phenomena represented in the case of a semi-speaker of Gaelic are (Dorian, 1977):

- absence of a stylistic option
- substitution of an analytic construction for a synthetic one
- analogical leveling

She notes that analogical leveling is the most characteristic practice of these speakers. She also observes that these speakers are pivotal figures in a process of language shift. It is clear that more attention needs to be given to these subjects. We will see later that, in the case of Basque, both in the homeland and in the United States, these speakers are growing in numbers and are becoming an important factor in the evolution of the language.

But this is not the only term that is ill-defined in sociolinguistic and ethnolinguistic research. There are also problems with the definition of certain terms such as 'mother tongue'. This term has been used according to monolingual assumptions, whereby one imagines a speaker acquiring first one language at home and then others, in a progressive, clean manner. This may be true of certain settings but not of all. Often one grows up speaking more than one language at home. One loses and gains proficiency in languages all throughout life. And there may be speakers who do not gain full proficiency in any language. By 'mother

tongue' we will understand here the language(s) in which one has full proficiency. This is still not a good definition because full proficiency has not been well defined either. In the same way that more research is needed to define what a semi-speaker is, there is a need also to do more research to define what a native speaker in a multilingual setting is.

Bilingualism as a socio-political situation can be divided into functional and symbolic bilingualism. An example of functional bilingualism is the case of warning signs written in English and Spanish in the United States, since they may save lives. In Montreal, every sign is written in French and English and this is an example of symbolic bilingualism (Mackey, 1983). Thus, bilingualism can be practiced de facto because of necessity, as in the United States without even having legislation concerning the use of language, like in the United States. The case of Montreal, however, exhibits uses of bilingualism that are just the result of the legal status of the languages and, hopefully, reflect the will of the speakers.

When bilingualism is taken as an individual, human phenomenon, there is a distinction to be made between a compound and a coordinate bilingual. A coordinate bilingual is someone who separates the registers and contexts of uses of the two languages. In this case, the two languages do not interfere very much with each other. A compound

bilingual will use the two languages in the same contexts and registers, often depending on the person(s) to whom s/he is speaking. This leads to interference, code-switching, borrowing, etc. But, some authors believe this distinction does not receive support from experimental data (Romaine 1995). Once again, more data are needed to fully characterize the different linguistic behaviors exhibited by speakers in multilingual settings.

Geopolitical divisions do not coincide with language divisions, not only because throughout the history of the establishment of political borders, linguistic status was rarely taken into account, but also because in the second half of the 20th century, ease of travel, readily-available means of communication, and political and economic upheaval/change have made people migrate. As in the case at hand with Basques in the U.S., we see communities abroad, diaspora states, or communities whose linguistic status is to be taken into account. We also need to note that in many ways, multilingual settings are not limited anymore to multilingual communities. Rapid worldwide communication and the ease of travel are exposing speakers to all kinds of multilingual settings never before experienced. In the case of communities with an important migrant component like the United States, this extends to environments as common as the work place. In the computer world, for example, there is

One of the speakers I interviewed, a Basque-American who is a professor at an American university, mentioned to me that on one occasion, years ago, she mentioned that she was Basque and spoke the language. The interlocutors were very surprised that she and her mother would be willing to use a peasant's language.

Some people are willing to face this type of pressure, while others, instead, just stop speaking what is characterized pejoratively in such a way. The avoidance of stigmatization can be a powerful motivating factor.

At some point, when the separation of functions breaks down, language choice serves the same functions as style shifting. (Gal, 1979) But this type of behavior does not have to be limited to a situation of language shift, as this is a common behavior in bilingual communities.

#### 4.3.2 Language Maintenance

Government-sanctioned Bilingualism in many countries exists in part to ensure the maintenance of more than one mother tongue in a certain territory, or to ensure the linguistic rights of individuals who are citizens of the same community but have different mother tongues.

As was mentioned above, in the history of migration to North America, maintenance of one's language alongside English has been

almost impossible. Some exceptions to this rule have occurred, such as the case of the Amish, some Hispanic communities, or Chinese-speaking communities. Enormous pressure is exerted upon the speakers and the outcome of this situation has been in most cases the loss of the mother tongue of the immigrant. One factor that has often helped speakers to maintain their language is religion and isolation. This is the case of the (German-speaking) Amish and some (e.g. Yiddish-speaking) Jewish communities. In the case of Basques, this is not a factor. Even though Basques are mostly Catholic, so are most of the Spanish-speaking immigrants with whom they interact constantly, as well as a good number of Anglosaxon Americans. Another factor that is often mentioned is physical appearance, or rather the clustering of like-speakers living in close proximity and interacting. This is the case, for example, of Asian immigrants. Being externally perceived as being "like-speakers" is not a feasible factor for Basque. Even though there are some physical characteristics that Basques can often detect, other populations are oblivious to them. Basque immigrants to the United States are thus perceived either as French, Spanish (most often Hispanic), or, simply, Americans. This fact has consequences on the external recognition of a community and forces Basques to constantly assert what they are.

acquisition of a second language, and to the inclusion in an environment where the opportunity (or social acceptance) for using the first language is severely reduced.

Hymes drew attention to a study by Hohenthal and McCorkle (1955) which contrasted two South American Indian groups in terms of retention of language (and identity). One group, the Fulnio of Brazil,

*have given up their lands several times during the last three centuries, moving in order to preserve their language and annual religious ceremony (to which proper use of the language is essential) as the basis of their identity* (Hymes, 1966:126).

For some communities, then, language maintenance takes precedence over maintaining their rights to a land. Each culture has different views on what culture entails and how language relates to their cultural integrity, religion, etc. This has to be taken into account when describing multilingual settings. Language maintenance is entirely in the hands of the community and associated with whatever functions the language is used for within communal life. Language, being a social and interactive skill, in the absence of a supportive social environment, becomes very vulnerable. As Cobarrubias (1983) notes, the status of language rights is in need of major clarification.

212

There are also individual factors that affect language maintenance. Some speakers are, for some reason, more apt to keep proficiency in their language than others under the same conditions.

#### 4.3.3 Language Attrition

Broadly defined, language attrition may refer to the loss of any language or any portion of a language by an individual or a speech community. It may refer to the declining use of mother tongue skills by those in bilingual situations or among ethnic minorities. It also refers to the deterioration of language skills in neurologically impaired patients. It may also describe the death of an entire language. Language attrition may also refer to the loss of language skills by those who have studied and then discontinued the use of a second language.

Little attention has been paid to the maintenance of these skills once attained and even less to what is believed to be the rapid decline in the functional command of these skills once they are no longer used.

Silva-Corvalán (1994), in her study of the linguistic behavior of bilinguals in Los Angeles, observes the following tendencies in the usage of the subordinate language:

213

*one-to-one mapping between underlying semantic structures and surface forms, with the goal of making messages easily retrievable for listeners.*

It seems, then, that the tendency in a process of language attrition is toward simplification and loss of forms. It should be noted, though, that the remaining forms may become semantically more complex, i.e. the forms left will have more meanings (Silva-Corvalán, 1994).

Foley (1988) argues that the fact that Ferguson's (1982) list of features of simplified language forms corresponds very closely to the universal features of pidgins in relation to their source languages suggests a close historical connection.

Simplification, though, can also be part of an elaborative process of language maintenance and even recuperation, in the sense that it can represent a stage in a process of language attrition, very similar to codeswitching (in the view of Silva Corvalán 1994, discussed above). That is, it can be part of the process of stabilizing a form of the language which (semi-) speakers can actually use.

Dorian (1978) argues, though, that language death cannot be equated with pidginization, given the data she obtained. But it is also clear that more studies of language extinction are needed in order to resolve the issue properly.

In general, the pattern found in studies of language shift and death is that formal stylistic options are reduced. Also, certain syntactic constructions accompanying formal styles become less frequent, e.g. subordination. Tsitsipis (1989), for example, documents the loss of features such as the gerundial construction in the Arvanitika variety of the Albanian spoken in Greece.

The studies of language attrition could be linked to studies in language acquisition since language acquisition studies can provide suggestions as to which linguistic features might be most vulnerable to loss and those which may be more robust.

While there are some universals in language acquisition, there may be great individual variation in either acquisition or attrition from such characteristics as differences in content and linguistic input, cognitive styles, gender differences, the likelihood of spontaneous rehearsal, and various speaker characteristics.

Also, loss rates seem to be language-specific (Pardee Lowe 1982).

One basic difficulty in the study of language attrition is that linguists cannot be sure of the level of linguistic competence at the onset of attrition. In the case of majority languages, being a native speaker is a criterion that is clearly generalized and can, thus, be used in composing

- Those morphological categories also marked in the stronger language will be maintained in the use of the weaker language.
- Those morphological distinctions which have a high functional load will be maintained.
- Those grammatical morphemes acquired earliest will be retained the longest and those acquired latest will be lost earliest.
- When a morpheme is overgeneralized, the morpheme that occurs frequently, is syllabic or free, and is least fused with other semantic features, will be the most likely candidate for overgeneralization.
- Greater number and variety of borrowed lexical items.
- There will be occasional use of innovated lexical items based either on the language s/he is speaking or on another language known by the speaker.
- Where elimination of a syntactic construction would result in information loss, that transformation will tend to be preserved, or the transformation will be eliminated and compensated for.
- There will be a preference for syntactic constructions using free morphemes where there is a choice between free and bound morphemes.
- Sporadic use of ungrammatical, non-native constructions reflecting analytic syntax.
- Where there are several competing forms for the same basic underlying meaning, one form will be overgeneralized and tend to replace the others.
- A highly regular form or construction will be chosen to replace an irregular form or construction (analogical leveling).
- Greater amount of paraphrasing and circumlocution.
- Tendency to choose a semantically related but partially inaccurate and inappropriate equivalent when using paraphrasing and circumlocution to compensate for linguistic gaps.
- There will be morphological and syntactic constructions based on the stronger language.

At the syntactic level:

- Use of a smaller number of syntactic devices.
- Preservation and overuse of syntactic constructions that more transparently reflect the underlying semantic and syntactic relations.
- Collapse of different surface structures into one.

sophisticated and late-learned aspects of language to be the first to disappear from the repertoires of second language speakers. A somewhat competing model was put forth in 1895 by Pitres, who held that the best-learned forms, whether learned early or late, should be the least vulnerable. Jakobson, in his work *Kindersprache, Aphasie, und allgemeine Lautgesetze* first published in 1941, adopted the regression hypothesis and claimed, *Aphasic losses reproduce in inverse order the sequence of acquisition in child language.*

The regression model does seem to hold for aphasics. It has not been shown, though, that it holds for language attrition and loss situations.

Berko-Gleason (1982) notes that it is very likely that the traditional linguistic subsystems (phonology, morphology, syntax, and vocabulary) may suffer differential loss in attrition.

Dorian (1982) observed that the imperfect speakers in her sample showed certain kinds of reductive phenomena in common in their Gaelic, regardless of acquisition history. For example: a greater or lesser use of analogically regularized allomorphs in place of irregular allomorphs; complete loss of morphemes that are already showing weakness in the fully fluent population's Gaelic; loss of "inventory," and also loss of vocabulary from both open and closed classes.

An important symptom is also the lack of purist reactions. In language attrition situations, the speaker is no longer aware of the errors and does not correct them. Even fluent speakers ultimately change their attitude and stop trying.

#### 4.4 The Case of Basque

As is known, Basque is a historically unclassified (or perhaps, unclassifiable) language isolate that is typologically very distant from its usual neighboring languages, Spanish and French. It is also very distant structurally from English.

It seems that at the beginning of the Christian era Basque-related dialects were spoken north and south of the Pyrenees and further east than at present time (Gardner 1998).

The first written Basque appears in the eleventh century under the form of a few words in a Latin manuscript. But the first literary texts date from the Middle Ages preserved by sixteenth century writers. There is also a printed text in Basque included by Rabelais in one of his books. But the first book, a collection of poems, was published in 1545. However, in the last twenty years the production of written Basque has exploded.

Most Basques, though, had little contact with written Basque. The shift from Basque to Spanish seems to have started around 1750 for reasons not yet well understood. The status of the Basque language continued to decline. The few attempts that were made to stop the decline of Basque came to a violent end when Franco occupied the Southern Basque Country in 1937. Franco's regime took every measure possible to prevent Basques from speaking their language or feeling any different from other Spaniards.

But this process was reversed already before the death of Franco in 1975 by a pro-Basque minority that was increasing in numbers. In the Northern Basque Country, however, Basque continues to decline vis-à-vis French.

A study published in 1995 by the Basque Government on the sociolinguistic situation of the Basque Country presents the following data: In 1991 there were 2,872,593 inhabitants in the Basque Country. Among those a little under 700,000 are believed to be speakers of Basque. The Basque Country (see page 227) is divided into three provinces in the Northern Basque Country, and four provinces south of the border. Among these, Bizkaia (17.5%) and Gipuzkoa (43.7%) contain the largest numbers of Basque speakers and contain considerable numbers of urban speakers. The Northern Basque Country 48.2%,

68.2%, and 53.7%) and Nafarra (10.2%) contain speakers that are, in general, more "rural". Araba (6.7%) is the province that contains the least amount of Basque speakers.

The political distinction that exists among the Basque Autonomous Community, Navarre and the Northern Basque Country is also reflected in ethnocultural and linguistic differences. Most of the Basque population considers itself Basque (64%), but a quarter of the population (24%) does not consider itself Basque. The remaining people are undecided.

The Basque Country south of the border that separates France from Spain is divided into two self-governing regions: The Basque Autonomous Community, which includes Gipuzkoa, Bizkaia, and Araba; and Navarre. The Northern Basque Country (also known as Iparralde) is located in French territory and divided into three provinces: Lapurdi, Baxenabarre, and Zuberoa.

There are no monolingual Basque among the population under the age of 50, but in The Basque Autonomous Community (from now on BAC), 34% of the population above the age of 64 is monolingual Basque.

On the negative side, this situation makes intergenerational transmission more difficult to implement. This standard variety, called 'Euskara Batua' (unified Basque) is composed mostly of features of Gipuzkoan. This is a variety that is not native to anyone, except for a subset of the new generations. In many cases grandparents have a very hard time understanding the grandchildren, but their grandchildren usually understand older speakers. An important sector of the older population rejects this variety.

At the outset, the choice of this device leads not only to the loss of dialectal varieties, but to the loss of forms in general in the language itself. The fact that Unified Basque is, in a certain way, a foreign language for everyone has produced consequences that are just now starting to be felt.

Onederra (1992) observes that one of the consequences of this and, more generally, of the process of resuscitation and standardization in general is the proliferation of semi-speakers. Among other examples, she cites the "overderivation" of certain words: *indartasuna* (indar-tasun-a 'strength-Nom-A'), instead of the existing word *indarra* (indar-a 'strength-A'). She also notes gender transfers from Spanish: *ihun akustika* ('acoustic wave'). Note that 'wave' is feminine in Spanish. But gender does not exist in adjectives in Basque, thus one would expect rather *ihun*

*akustiko*. The change in final vowel in the borrowed adjective seems due to a gender perception that comes from Spanish or French.

#### 4.4.2 Basque in the Diaspora

We will dwell on the question of Basques in the United States in the next chapter. But, in general, Basques have migrated throughout the last four centuries. The types of migrations have been different because the causes have been different: conquest, economic survival, war, etc.

The Basque Government has created, in the last few years, a special department dedicated to the relations with Basques in the diaspora. The need has been felt because of the pressure from these communities to be recognized as Basque and because of a lesson learned in the last war. During the 1936 war, caused by the fascist coup d'état, the Basque Government in exile received important support from Basques established abroad.

Basques have migrated all over the world, but especially to the Americas. There are very important Basque communities in Argentina, Chile, and Mexico, as well as in the United States. As an example of the strength of some of these communities, it is worth mentioning that the University of Nevada at Reno has a Ph.D. program in Basque Studies.

In the next chapter I apply both the notions discussed in this chapter and the observations and conclusions reached in Chapter 3 to a set of speakers.

## CHAPTER 5

### RESULTS AND DISCUSSION

#### 5.1 Introduction

In this chapter I will present and analyze the data I collected in my interviews. I did not use statistical methods. Background linguistic data, i.e. a baseline, on the speakers was not available and no previous study of this type had been done on Basque speakers in the United States. Instead, I used in-depth interviewing and as a baseline I used two native fully fluent speakers, a Bizkaian and a Gipuzkoan (see Appendix A).

This method limits the number of subjects one can investigate. For obvious reasons, the interviews themselves, as well as the analysis of the data take a long time. Using this type of interviewing technique, one loses sight of the larger picture but gains a deeper understanding of the situation. Most of the interviews were done in Basque. I, myself, am a speaker of Basque and, despite the fact that I have never received

spreads between the 16th and 17th centuries, and another one that corresponds to the 19th and 20th centuries

The earlier wave was one of administrators, conquistadors, and founders. The later wave represents more an "exile" due to excess in population and lack of resources in the Basque Country. The recent wave of Basque exile, the one due to the Spanish Civil War (1936-39), should also be mentioned here.

There were a group of Basques on Columbus' first trip (most of them aboard the Santa María and a few on the Niña). The second group went with Hernán Cortes. This group, composed of mercenaries, was augmented by a later influx of scholars, colonists, and religious people. Basques were establishing here an important tradition of sailors and travelers (Elcano, for example, was the first person to finish a trip around the world in 1552. As Magellan's first officer, he took over when the captain died in the Mariana Islands). All of these travelers were seeking success and escaping the limitations of their former lives. A case in point is Catalina de Erauso, who spent most of her life as a male soldier in the Americas, escaping the limitations of traditional female roles in Europe.

The 18th century witnessed commercial relationships between America and the Basques. Under Philip V, the Real Compañía Guipuzcoana de Caracas was created in 1728 for commerce.

240

In the 19th century, Basque immigration to the Americas was driven by the excess population subsisting on commerce, industry and fishing, the decline in agricultural productivity, wars, and the land inheritance system.

Around 1850, the maritime industry began developing in Argentina. Towards the end of the 19th century, massive immigration to Argentina ensued, raising inflation and pushing Basque immigrants to other lands, such as California. The occupations engaged in by Basques began to diversify as well.

Basque associations started proliferating in 1876. Basque publications started to appear abroad in 1878.

Basque nationalism of the last two decades of the 19th century profoundly influenced the thinking of the Basques in the Americas.

In regard to the Diaspora of the 20th century, most of Basque immigration has taken place after 1936. The immigration dating from the first decades of the century was prompted by economic hardship, mostly resulting from World War I.

Starting in 1903, new Basque nationalist organizations overseas were consolidated in three centers: New York, Mexico City, and Buenos Aires. There was a new component in this immigration, which consisted of

241

addition, New York has always attracted many foreigners in general due to job opportunities and intellectual activities.

### 5.1.2 Basques in the New York Area

The goal of this dissertation is to study, once it has been established that nominalizations are complex syntactic operations, the use of these nominalizations made by Basques in the New York area, both Basque Americans and Basque immigrants. The fundamental assumption is that this community exemplifies semi-speakers, as well as full speakers, of Basque, who are far removed from linguistic environments in which their language abilities may or may not continue to survive and grow. Therefore, this study is not concerned with language attrition that occurs in a large population in its native land. Rather, it is concerned with language attrition in a "diaspora" situation.

In 1980 Basques in the United States were allowed to enter their ethnic origin in the Census. The numbers are as follows (White 1993):

State	French Basques	Spanish Basques	Basques Undifferentiated		Total Basques
			(0)	(46)	
Alabama	24	(36)	44	(0)	14 (82)
Alaska	37	(10)	38	(33)	170 (105)
Arizona	53	(152)	298	(199)	965 (1,100)
Arkansas	20	(34)	21	(0)	63 (73)
California	3,387	(3,619)	3,508	(3,813)	12,227 (15,530)
Colorado	148	(341)	110	(168)	679 (446)
Connecticut	22	(36)	64	(64)	233 (120)
Delaware	0	(18)	7	(0)	6 (3)
Dist.of Columbia	0	(22)	16	(12)	21 (29)
Florida	117	(201)	334	(315)	738 (343)
Georgia	11	(87)	27	(59)	90 (77)
Hawaii	19	(10)	29	(4)	121 (55)
Idaho	166	(221)	353	(600)	5,068 (3,511)
Illinois	49	(422)	75	(66)	321 (165)
Indiana	55	(94)	0	(48)	135 (18)
Iowa	20	(260)	8	(24)	31 (40)
Kansas	10	(92)	24	(18)	36 (50)
Kentucky	11	(81)	15	(15)	68 (36)
Louisiana	73	(133)	38	(57)	115 (65)
Maine	2	(22)	21	(0)	13 (28)
Maryland	60	(51)	45	(48)	163 (148)
Massachusetts	37	(34)	73	(80)	227 (187)
Michigan	7	(145)	47	(28)	162 (158)
Minnesota	24	(110)	15	(8)	91 (102)
Mississippi	4	(7)	0	(2)	24 (20)
Missouri	27	(164)	10	(18)	114 (61)
Montana	66	(116)	46	(6)	357 (268)

*(to be continued)*

Table 2. Basque Population of the United States - 1990 and (1980)

### 5.2.2 Age

In the Basque community in the New York area all ages are represented, since they have formed families. One age group that is quite small is that of elderly individuals. Most often Basques will go back to the Basque Country, once retired, leaving their children behind. Actually, there are procedures put in place by the Basque Government to make this "re-incorporation" process easier. Despite the fact that they are American, these individuals are being recognized as Basques in terms of the law and the transfer of their pensions is being facilitated.

Among my subjects there were two girls of age 10 and 11, and a teenage boy. The adults were mostly male ranging from 26 to 82 years of age. The females ranged from 24 to 73 years old.

### 5.2.3 Education

Most informants have attained the high school level in education. Among them I can count one PhD, and a couple of Bachelor's Degrees. They work mostly in construction, the older ones coming from sheep herding backgrounds. They have certainly never been educated in Basque (except for two cases), and, therefore, the Basque they speak comes, essentially, from their parents and their interaction with other Basques, both in the US and in Europe.

### 5.2.4 Speaking Abilities

Among the subjects interviewed, 27 are considered to be speakers of the language. As was mentioned in Chapter 4, semi-speakers exhibit conversational devices that make them appear as full speakers in the community. Once these speakers are deprived of those devices and are tested and observed, we see that their speaking abilities have serious shortcomings. After analysis of the data it became clear that except for three or four of the speakers, most of the subjects are to be considered semi-speakers or speakers undergoing a language attrition process.

Most of the speakers use their dialectal variety. Even though most of them are literate, only two read or write in Basque, and mostly only correspondence. And all of them use the language only occasionally. In one case, it was claimed that Basque was the only language spoken at home. This fact could not be verified and I questioned the children of the speaker on this issue, they contradicted their mother. The mother did always address the children in Basque, but the children did not always respond in that language. In addition, the frequent presence of non-speakers in the home made it impossible for the language to be spoken constantly. It is true, however, that this is one of the few families where the children are able to hold a conversation in Basque. Most of these speakers also speak Spanish.

Most of the speakers I interviewed are trilingual, Basque-Spanish-English. The Basque Americans are usually bilingual, Basque-English, and have some knowledge of Spanish.

### 5.3 The Data

For the collection of the data I used two types of questionnaires and interviews. One of the questionnaires contained a set of 72 sentences to be translated into Basque, from Spanish or 65 from English. The other questionnaire contained a set of 91 sentences in Basque containing structures to be tested. The informants were asked to approve or disapprove of the sentences and, in any case, give the version they would use in their own "dialect".

One aspect that seemed at the beginning to be a difficulty turned out to be an advantage. I used Unified Basque in my examples, as this is the standard version of Basque that was created in the process of normalization of Basque in the Basque Country. This version of Basque is quite similar to the Gipuzkoan dialect that I speak. Some of the lexical items were unknown to the informants, who are almost all Bizkaian. This turned out to be an advantage, as most of them assumed this was my dialect. The reaction of the informants almost unanimously was to let me know how they would say it in their dialect. This allowed me to collect more items.

252

I also held a conversation-type interview with each informant. This was a requirement for several reasons. I wanted to collect information on language attitudes, ethnic identity and national identity. Not only language attitudes, but also ethnic and national identity sentiments are fundamental for determining the survival of a language. When possible – in most of the cases as it turned out- the interviews were done in Basque, allowing me, again, to collect more linguistic data. Furthermore, this data was more natural, since the informants got involved in the conversation and were less careful with their way of speaking. All these interviews were tape-recorded.

Unfortunately not many nominalizations showed up in these sections. But clues about their level of proficiency did appear.

In many cases, I had to read and write the sentences for them. On the one hand, several of them were quite afraid of reading and writing in Basque. They wouldn't have made it through the questionnaire if we had not filled them out together orally. On the other hand, since they did not know all the lexical items, it was imperative for me to be there and assist the informants.

I used two sets of baselines, one in the Gipuzkoan dialect, and one in Bizkaian, the latter being the most common one among the

253

(124) \*Zuk harakinari ni saldu n(a)iozu  
you-E butcher-D I-A sell-prf Aux/1sgA.3sgD.2sgE

(125) Zuk ni saldu nazu karnizeruri  
you-E I-A sell-prf Aux/1sgA.sgE butcher-D  
'You sold me to the butcher'

Sentence (125) exemplifies the utterances produced by the speakers. They had no problem with the sentence. This could be due to the fact that they are having difficulties with the auxiliary system, and, thus, have lost the restrictions. It could also simply be due to the interviewing context, whereby the speaker is undergoing an exercise, where s/he is asked to produce utterances.

Then, of course, they did not have any problem with allowing all the arguments in the nominalization. They allowed everything and, since they did not have to build the auxiliary, they had no problem with it.

In general, the fact that the nominalization requires a determiner or that it needs to be case-marked is respected. Some speakers, though, did have a tendency to leave the *-te* nominalization indefinite in certain contexts. I do not have an explanation for this.

(126) a. Zuk ematea nahi dut [g]  
Zuk berari **emotea** nahi dot [b]  
You-E s/he-D give-nom-A want Aux/3sgA.1sgE

b. nai dot **emoti**  
want Aux/3sgA.1sgE give-nom-bare  
'I want you to give it to him/her'

In (126b) the speaker should have produced *emotie*.

The temporal constraints are recognized. All speakers judged example (127) ungrammatical. In (127) the subordinate clause is in the future tense. And most of them correctly produced (128). Also, no one used a nominalization in environments such as 'he told us that he saw a cow flying'. They all seem to know that the nominalization cannot refer to an accomplished past and produced subordinate clauses instead.

(127) \*Batzarretan **hitegingo dudala** debekatu  
meeting-pl-loc speak-fut Aux/3sgA.1sgE-comp forbid-prf  
*didate*  
Aux/1sgD.3sgA.3plE

(128) Batzarretan **hitzegtea** debekatu didate [g]  
meeting-pl-loc speak-nom-A forbid-prf Aux/1sgD.3sgA.3plE  
'They forbade us to speak at meetings'

Also, except for two cases, they knew they could not have a sentence as a subject of a transitive verb. It is to be noted, though, that when they create a derived nominal or build a *-te* nominalization in this type of environment, they get confused in the case marking.

#### 5.4.3 Distribution of -te Nominalizations and Derived Nominals

In contexts such as subject of the matrix clause or as complements to verbs like *nahi* 'want,' *ahaztu* 'forget,' *lortu* 'succeed,' or to structures with predicatives such as *erraza* 'easy,' group A speakers hesitate and use non-nominalized gerunds, infinitives, or subordinate clauses. They tend, correctly, to use *-te* nominalizations with verbs like *gorrotu* 'hate' and *arritu* 'surprise'.

(132) a. ***Idaztea***   *ahaztu*   *zitzaidan* [g]  
 write-nom-A forget-prf Aux/past.1sgD.3sgA  
*Aztu ein jaten **idaztea*** [b]  
 b. *Agestu*   *naz*   *eskribiduten*  
 forget-prf Aux/1sgA write-ger  
*Astu   egin   zait*   ***eskribidu***  
 forget-prf do-bare Aux/1sgD.3sgA write-inf  
*Astu   egin   dast*   *selan **escribidu***  
 forget-prf do-bare Aux/1sgD.3sgA how write-inf  
 'I forgot to write'

In (132b), *eskribiduten* is a gerund, and *eskribidu* and *escribidu* are infinitives.

(133) a. *Aita*   *gaixorik egotea*   *sentitzen dut* [g]  
 father-A ill   be-nom-A feel-ger Aux/3sgA.1sgE  
*Sentituden dot **aita gaixorik egotea*** [b]  
 'I regret father's being ill'  
 b. *Sentituden dot*   *aita*   *gaixorik dauela*  
 feel-ger Aux/3sgA.1sgE father-A ill be/3sgA-comp  
 'I regret that father is ill'  
(134) a. ***Idaztea***   *ahaztu*   *zitzaidan* [g]  
 write-nom-A forget-prf Aux/past.1sgD.3sgA  
*Aztu ein jaten **idaztea*** [b]  
 'I forgot (how) to write'  
 b. *Astu*   *einjat*   ***irakurri***  
 forget-prf Aux/past.1sgD.3sgA read-inf  
 'I forgot to read'??  
(135) a. ***Batzarretan hitzegitea***   *debekatu*  
 meeting-pl-loc speak-nom-A forbid-prf  
 didate  
*Aux/1sgD.3sgA.3plE*  
 'They forbade me to speak at meetings'

speakers preferred instead to make longer sentences or use subordinate clauses, even at the expense of changing the meaning of the sentence.

(138) *Aitaren berebilaren erosketa oso ona izan*  
father-G car-G purchase-A very good be-prf

*da*

*is*

'Father's purchase of the car has been good'

(139) *Ordu onin erosil deu aita*  
hour good-loc buy-prf Aux/3sgA.3sgE father-A  
*cotxie postu nas*  
car-A become happy Aux/1sgA

'At a good time did father buy a car, I'm happy about it'

When dealing with adjective modification of a nominalization with *-te* and a demonstrative, most of the speakers (most from group A), replaced the structure with a sentential structure and adverbial modification, or simply changed the sentence.

(140) a. *Aitaren ibiltze azkar hori gustatzen*  
father-G walk-nom-bare fast that please-ger  
*zat*  
Aux/1sgD.3sgA  
'I like father's that fast(adj) walking'

b. *...aitaren azkar ibiltzea*

father-G fast walk-nom-A

'...father's fast(adv) walking'

(141) a. *Aitonaren lan egite isil hora*  
grandfather-G work do-nom-bare silent that-A  
*gustatzen zitzaidan*

please-ger Aux/past.1sgD.3sgA

I liked that silent working of grandfather's

b. *Nire aititen biarra egiten bera bakarrik*  
my grandfather-G work-A do-ger himself alone  
*gusten dozte*  
please-ger Aux/?

I like my grandfather's working (gerund) him by himself

(142) *Keparen hizketa lasai hori maite dut*  
Kepa-G speech calm that-A love Aux/3sgA.1sgE  
'?I like that calm talking of Peter's'

(143) *Keparen lasai hitzegitea maite dut*  
Kepa-G calm speak-nom-A love Aux/3sgA.1sgE  
'I like Kepa's calm talking'

(144) *Keparen hitzegiteko lasaitasun hori...*  
Kepa-G speak-nom-G calm(N) that-A  
'Kepa's calm for talking...'

(146) *Lehendakariak presoen askapena eta*  
 President-E prisoner-pl-G liberation and  
*suetena* negoziaztea eskatu  
 cease-fire negotiate-nom-A request-prf  
*ditu/du*  
 Aux/3plA-3sgE/3sgA.3sgE

(147) *Lendakariak zuzendatu zun suaren*  
 President-E request-prf Aux/past.3sgA.3sgE fire-G  
*itzaltzea eta presoen askatasuna*  
 turn off-nom-A and prisoner-pl-G liberation-A  
 'The president requested the liberation of the prisoners and  
 the negotiation of a cease-fire'  
*Lendakariak demandaten do prisoneruen*  
 president-E request-ger Aux/3sgA.3sgE prisoner-G  
*liberazinoa eta gelditu tiruak*  
 liberation-A and stop-inf shot/A.pl  
 'The president demands the liberation of the prisoner and  
 to stop the shooting'

When the speaker was prompted to produce or to evaluate a sentence such as (146), at best s/he produced (147). Note that in (146) we have the possibility of the plural auxiliary (*ditu*) or the singular one (*du*). Not only speakers in my study used the singular (*zun* and *do*) -if they were able to produce the sentence at all-, but they also corrected

the sentences presented to them in Basque that showed grammatical and ungrammatical pluralization. They just defaulted to the singular. The singular is acceptable, but it is noticeable that, with only two exceptions, they did not even contemplate the possibility of the plural. Note, also, that in the second utterance under (147), the speaker is coordinating a derived nominal (*liberazinoa*) with an infinitival (*gelditu*).

When using a temporal modifier with a nominalization that already has a modifier in the genitive, speakers exhibited confusion. The temporal modifier should take the marking *-ko*, which is one type of genitive, and there is already another modifier with another genitive, *-ren*. The speakers mostly tried to avoid the situation or to use other case markings.

(148)a. *Aitaren gaurko etorrera*  
 father-G today-G arrival-A  
*Aitaren gaurko etortzea*  
 father-G today-G come-nom-A  
*\*Gaurko aitaren etortzea*  
 today-G father-G come-nom-A  
 'Father's arrival of today'

responded. I also had one speaker that produced the ungrammatical string.

(151) *Gurasoen Ondarrura etortzea atzeratu*  
 Parent-pl-G Ondarru-adl come-nom-A push back-prf  
*dute*  
 Aux/3sgA.3plE  
*\*Gurasoen Ondarrura etorrera atzeratu*  
 parent/G.pl Ondarru-adl arrival-A postpone-prf  
*dute*  
 Aux/3sgA.3plE  
 'They postponed the arrival of the parents to Ondarru'

Note that the structure is grammatical when we use a *-te* nominalization (*etortzea* 'coming') and ungrammatical when we use a derived nominal (*etorrera* 'arrival').

In regard to wh-extraction restrictions, they avoid the structure that would make them choose. I could not elicit it. They kept using *-te* nominalizations, which do not require extraposition. The relevant contrast is presented in (152).

(152) *Noren usteza da Ainhoak garagardoa*  
 Who-G belief-instr Aux/3sgA Ainhoa-E beer-A  
*ekartzea beharrezkoa?*  
 bring-nom-A necessary  
*Noren usteza da beharrezkoa Ainhoak*  
 Who-G belief-instr Aux/3sgA necessary Ainhoa-E  
*garagardoa ekartzea?*  
 beer-A bring-nom-A  
*Noren usteza da beharrezkoa Ainhoak*  
 Who-G belief-instr Aux/3sgA necessary Ainhoa-E  
*garagardoa ekar dezala?*  
 beer-A bring-bare Aux/subj.3sgA.3sgE.comp  
*\*Noren usteza da Ainhoak*  
 Who-G belief-instr Aux/3sgA necessary Ainhoa-E  
*garagardoa ekar dezala beharrezkoa?*  
 beer-A bring Aux/subj.3sgA.3sgE.comp necessary  
 'In whose opinion is it necessary that Ainhoa bring beer?'  
 It could be that I could not elicit it because they just do not use subjunctives (*ekar dezala* 'bring'), yet I am quite sure some of them actually do. It could be that they were aware of the constraint, and the difficulty it poses, and avoided it.

Only speakers in group B accepted the compound formation with derived nominals.

Speakers in group B reacted to the ungrammaticality of a *-te* nominal with a genitive subject and a dative argument.

(156) *\*Andoniren haurrei ipuin kontatzea*

Andoni-G child/D.pl story tell-nom-A

'Andoni's telling stories to the children'

The observations above point to speakers with a subordination and nominalization system that has broken down. Even though, in my departing hypothesis I expected the speakers not to use *-te* nominalization, my research shows different results.

The speakers in group A are, in my opinion semi-speakers (according to Dorian's (1982) definition).

*They can be distinguished at the lower levels of skill from people who know only words and a few fixed phrases by their ability to MANIPULATE words and form sentences. They can be distinguished at the upper levels of skill from the youngest of the fully fluent speakers by the presence of deviations in their dialect which are generally recognized by the rest of the community as "mistakes."* (p. 32)

They overgeneralize the use of the *-te* nominalizations in the contexts where it would be more appropriate to use derived nominals.

They tend to use tensed clauses where they should use *-te* nominalizations. They have lost most of the criteria that would allow them to choose among a gerund, an infinitival, and a *-te* nominalization. They have lost the ability to use and distinguish the uses of a *-te* nominalization or a derived nominal with a demonstrative. They have lost the restrictions on coordination of nominalizations and derived nominals. They have lost the restrictions on pluralization of those coordinations. In addition, they have lost a good part of the auxiliary system.

#### 5.4.4 Sentential versus Nominal

There is a general tendency for all speakers interviewed to lean toward sentential structures rather than nominal ones. This is shown not only in their preference to construct the nominalizations with the arguments showing the same case marking that they would exhibit in a sentence, but also by their preference to use adverbial modification, rather than adjectival. Another aspect of their linguistic behavior relevant here is the fact, also mentioned above, that they dislike the use of demonstratives with nominalizations, and the fact that they do not produce many derived nominals.

#### 5.4.6 Complexity

When the sentences to be translated became longer and required more vocabulary, most of the speakers gave up. Only speakers in group B and one speaker from group A attempted the exercise. When the same sentences were presented in Basque, the attitude was similar. Many of the speakers complained about the use of strange words and about the fact that they did not know how to read and write in Basque.

This adds evidence to the argument that these speakers are in a process of language attrition. They build and prefer shorter sentences, and they lack vocabulary.

#### 5.4.7 Code Switching

Code switching occurred during the conversations concerning ethnicity. These speakers seem to practice code switching less frequently than do the Spanish speakers studied in Silva-Corvalán (1994), or Hispanics in New York with whom I have spoken (impressionistically speaking). Hispanics seem to exhibit a very complex and sophisticated system of code switching. In the case of my speakers, code-switching occurred only occasionally. More than anything else, the Basque speakers exhibit a high rate of lexical borrowing and when, during the

conversation, they get tired or do not have the words needed, they just switch to one of the other languages.

The code switching that occurs with most of these Basques is a three-way code switching, involving Basque, Spanish, and English, even in the case of Basque Americans. This is not to be generalized to all Basque communities in the United States. Each community has developed different patterns, due to their different histories and different origins. The communities in the western United States seem to be quite different. Some examples of codeswitching include:

**Luis y ella hablan en vasco, pero solo lleva 7 años.** Ni americanoa, bai, euskalduna jaio naz baino nazionalidadea. Zergatik euskalduna? Ez dakit....

Hau jaio zanian americano ziudadano zan. **Andoni was an accountant, I was secretary with Aguirre ta gero Social Security representative.** Ni bai, ni eskolan egiten duzenan bost urte edo....

Badakije euskaldune zer dan. Ni euskalduna, bai. Amen Amerikan ez dakite diferentie dala, **genetic....Nik ez dakot konfliktorik, bijak, the best of two worlds.**

#### 5.4.8 Attitudes

While the linguistic data did not allow for a categorization of the speakers along the lines of place of birth and education, in regard to attitudes toward ethnicity, the split is quite clear. The attitudes of subjects raised in the Basque Country are very different from those of Basque Americans.

Basque Americans, whom I define as those born and raised in the U.S., not those who have an American passport, feel they are both Basque and American.

I think I'm American Basque. Although I am born in the US, my heritage is Basque. The country in which I was born is the US, therefore I'm an American. It means that I have more privileges than people in most parts of the world. But it's different from heritage.  
*[female, 48 years old]*

Soy americana, pero me siento más de allí. Yo, a veces, cuando voy allí digo que soy americana, o digo que soy vasca americana y aquí digo que soy vasca. *[female, 24 years old]*

*'I am American, but I feel I am more from there. Sometimes, when I go there I say that I am American, or I say that I am Basque American and here I say that I am Basque.'*

Both parents speak Basque to each other, to me too. So, I know Basque, superficially, but I do. We use the three languages at home. I'm proud, it's a unique language and culture that I'm proud to be a part of. *[male, 29 years old]*

When things were unclear and I asked the subjects if they were more like my husband (from Ohio) or more like me (Basque born and raised in the Basque Country), they all agreed that they were more like me. The younger subjects reacted with expressions such as, 'No way! Like you!', even if they did not know a word of Basque, their Spanish was defective, and their English was native. Their perception of themselves appears quite detached from their language abilities.

Some Basque Americans argue that they would rather be American, in nationality, given that Spain has always been cruel to the Basques.

Guk langile oso onak gara. Alemaniarrak ere oso langitxiak dira ere baino...Desberdintasuna da nondik istorri gara. *[male, 30 years old]*

'We are very good workers. Germans are also very work-oriented, but... The difference is where we come from.'

Euskalduna izatea importantiena da. Ni jaio naiz euskalduna ta...  
*[female, 48 years old]*

'To be Basque is the most important thing. I was born Basque, so...'

Mucha cultura, jo, cultura aquí no hay. Yo veo las raíces, roots you know, las raíces, no sé cómo explicar. *[female, 24 years old]*

'A lot of culture, here there is no culture. I see the roots, 'roots you know,' I don't know how to explain.'

*It's very hard for us to relinquish a thought without being sarcastic or sardonic. There are also physical attributes also, the noses, the cheeks, the ears...The hygiene, people in the Basque Country are less clean than here. [male, 30 years old]*

Temperamentue geiena independiente, ezin dozu sakatu asko oni baino baino konsegiduko dozu nai dozuna...por ejemplo euskaldunek aintzinez okin dabie euren kalendariojue, euren matematikie,..euskaldunek esplikau leuzen nundik sartun den Colon amen, ze beran kartografue izen da euskaldune ta ekarri tzoan. *[couple, 73 and 74 years old]*

'Mostly the personality is independent, you cannot push them a lot, but you get what you want from them...for example from a long time ago Basques had their calendar, their mathematics...Basques explained how Columbus got here because his cartographer was Basque and he brought him here.'

Even though, as can be seen, racist feelings and 'romantic' ideas arise here, in general, all subjects agreed that to be Basque was to be born Basque. This is an issue that is very unclear in the Basque Country. Children that are born in the Basque Country are Basque in the eyes of the law. There is still a great deal of debate over this. The reason for this is the racial issue.

Where all the subjects come together, also, is in regard to the maintenance of the language and the culture. They all feel it should be done, both Basque Americans and Old World Basques, for the sake of

These speakers exhibit all the symptoms of a community undergoing language attrition. As Dorian (1981) mentions in her study of Gaelic, these speakers live undetected in the community thanks to the survival devices they have developed. I felt, like Dorian did, that I was revealing a dirty family secret by eliciting the data. Without any further analysis the speakers became aware, as we went through the exercise, of the fact that they knew less Basque than they thought.

The avoidance mechanisms varied. The speakers sometimes refused to undergo parts of the exercise, while others accused me of being sadistic and asking for an impossible task. One speaker even refused to translate some of the sentences because reading, she said, made her dizzy. I read the sentences for her and she was still having problems producing the utterance.

This is a delicate issue. Immigrants are torn between their past, the warmth of their childhood and youth, and their present, with its new loves, friendships and languages.

More linguistic research needs to be done in this area. This type of work sends a warning sign about a community and their language. This type of warning sign may allow us to correct the problem in time. In addition, the way in which semi-speakers deal with linguistic structures

and linguistic complexity provides data that may contribute to an effective analysis and interpretation of linguistic structures.

In regard to the ethnic identity attitudes exhibited here, Douglass (1996) seems to believe that Basque communities are in decline in the United States. He may be correct, I will not try to predict unannounced deaths, but the energy to survive is certainly there.

As far as their use of nominalizations is concerned, speakers overgeneralized the use of verbal *-te* nominalizations. This behavior probably reflects the fact that verbal *-te* nominalizations are syntactically transparent. This type of nominalization is transparent in the sense that case marking in this nominalization is the same as in a finite clause. The speaker presumably learns the finite structure earlier, and this makes the nominalization very "easy" as to generation.

I did not find any evidence pointing to an influence from English or Spanish in the distribution of *-te* nominalizations among the speakers. Rather, the linguistic behavior exhibited seems due exclusively to a process of language attrition and the simplificatory pressures in such a situation.

The data collected for Basque show that, despite my initial hypothesis that nominalizations are a complex structure and, thus, should be lost early in the process, they are maintained. However, they are not maintained wholly intact, since the speakers overgeneralize one subtype of these nominalizations and have lost a great many of the restrictions that apply to these structures in Standard Basque, thereby simplifying them. It appears that the syntactically most transparent nominalizations are kept and generalized.

However, what is most crucial here are not the details of the syntactic derivation of nominalizations in Basque rather the ethnolinguistic setting for the use of the nominalizations. The purely linguistic analysis is not irrelevant, to be sure, but it is not all there is. One should note that even a semi-speaker who gets Basque all "wrong", according to the standards of speakers considered fully proficient, is still using Basque "successfully" in a certain sense, in that they are able to fulfill their communicative needs within the context of their diasporic Basque compatriots. They are successfully negotiating "Basqueness" with their language use.

In this type of research linguists end up being like anthropologists who necessarily have to reflect on their methods constantly because their being a participant-observer, necessarily alters what is being observed.

Thus the linguist - any linguist, but especially an anthropological linguist - must be self-reflective and introspective about methods and results. As an anthropological linguist, therefore, I confront the purely linguistic issues and methods I have employed, and must wonder indeed how relevant the details of syntactic derivation are, e.g. whether an empty subject position is pro or PRO, whether a head can be category neutral, where the derivational component applies, etc.

While it may seem like an indictment of linguistic practice to say this, perhaps what is relevant to the speaker is just the availability of a certain derivational device, not whether this device is formalized as syntactic or lexical in nature. In this way, one must wonder how seriously we need to take Chomsky's assertion that E-language is irrelevant.

## APPENDIX A

### SPANISH SENTENCES GIVEN TO SUBJECTS & BASELINE

1. El haber matado a Calvo en la Cárcel está mal

Calvo kalabozetaan akatza larria da

calvo espetxeen hil izana gaizki dago [g]

mikel kartzelan hiltzea txarto ereizten jat [b]

2. Lamento que el padre esté enfermo

Sentitzen dut aita gaiko izatea

aita gaixorik egotea sentitzen dut [g]

sentiduten dot aita gaizorik egotea [b]

3. Leer este libro es fácil

Erraza da liburu hau irakurtzea

liburu hau irakurtzea erraza da [g]

erreza da hau liburua leidutea [b]

4. El padre me deja hablar

Aitak itzegitea utzi nau

aitak hitzegiten uzten dit [g]

aitak berba egiten izten nau [b]

5. La madre no ha conseguido que el niño coma

Amak ez du lortu umeak jatea

amak ez du lortu haurrak jatea (ambiguo)[g]

(ama saiatu da baina, haurrak ez dio/du jan)[g]

amak ez dau lortu umiak jatea [b]

6. Se me olvidó escribir

Ahaztu zait idaztea

idaztea ahaztu zitzaidan [g]

aztu ein jaten idaztea [b]

7. Nada más entrar Gorka en la habitación nos fuimos

Gorka gelan sartzean, alde egin genuen

gorka gelan sartu bezalaxe, alde egin genuen [g]

gorka kuartoan sartu ta gero jun ein ginian [b]

8. Pensaba cenar

Afaltzea pentsatu nuen

afaltzea pentsatzen nuen [g]

afaltzea pentzeten neban [b]

9. Me sorprende que el padre esté tan bien

Aita hain ondo izateak arritu nau

aita hain ondo ikustea harritu egiten nau [g]

sorprendidu egiten nau aita ain ondo egotea [b]

10. La compra de la madre ha sido buena

Amaren erosketa ona izan da

amaren erosketa ona izan da [g]

amak egin daben erosketa ona izan da [b]

11. La compra de los libros y los vestidos

Liburuen eta jantzien erosketa

liburu eta jantzien (soinekoen) erosketa [g]

liburuen eta erropean erosketa [b]

12. La compra del coche por parte del padre ha sido muy buena

Aitaren berebilaren erosketa oso ona izan da

aitak egin duen kotxe erosketa oso ona izan da [g]

aitak ein daben erosketa kotxeagaz ona izan da [b]

13. Me gustaba el silencioso trabajar del abuelo

Aitonaren lan egite isil hura gustatzen zitzaidan

27. La compra de tus libros de aquella tienda

denda hartako zure liburuen erosketa [g]

hango dendako zure liburuen erosketa [b]

28. La lectura de los estudiantes de libros en casa

Ikasleen etxeko lan irakurtzea

\*Ikasleen etxean liburu irakurtzea

? ikasleen liburu(k) irakurketa etxean [g]

ikasleak etxean liburuak irakurtzea [b]

29. Decidi estar en silencio

Isilik egotea erabaki dut

isilik egotea erabaki nuen [g]

izilean egotea erabaki neban [b]

30. Sabemos cómo fastidiar a la gente

Badakigu jendeari adarra jotzen

jendea nola izorratu badakigu [g]

badakigu zelan jentia fastidija [b]

31. No me atreví a decir nada

Ez nintzen ausartu ezer esatera/esaten

eznintzen ezer esatea/esaten ausartu [g]

ez naz atrebiduten ezer esaten [b]

32. Nos dijo haber visto una vaca volando

Idi bat hegan ikusi duela esan digu

\*Idi bat hegan ikustea esan digu

behi bat hegan ikusi zuela esan zigun [g]

bei bat bolando ikusi ebala esan euskun [b]

33. Me prohibieron hablar en las reuniones

\*Batzarretan hitzegingo dudala debekatu didate

Batzarretan hitzegitea debekatu didate

bileretan hitzegin nezan debekatu zidaten [g]

batzarretan berba egitea galazotu egin eusten [b]

304

34. Me prohibieron que hablase en las reuniones

bileretan hitzegin nezan debekatu zidaten [g]

batzarretan berba egitea galazotu egin eusten [b]

35. Me alegré de que fueras a América

Zu Ameriketara joateak poz eman zidan

poztu egin nintzen zu amerikara joatearekin/joateaz [g]

poztu nintzan zu ameriketara juteagaz [b]

36. No me alegré de que fueras a América

Zure Ameriketara joateak ez dit pozik ematen

ez nintzen poztu amerikara joatearekin/joateaz [g]

[ez nuen poz handirik....(jatorrago)] [g]

ez nintzan poztu zu ameriketara juteagaz [b]

37. Me sorprende que María llegue tarde

Miren berandu heltzea harrigarria da

maria berandu iristea harritu egin nau [g]

sorprendidu egiten nau miren berandu allegatea [b]

38. Me molesta que María no pueda estar callada

Gorrotu dut Miren ixildu ezin izatea

mariaren isilik ezin egon horrek molestatu egiten dit [g]

miren izilik ezin egotea molesta egiten nau [b]

39. Odio ese continuo hablar sin saber de Kepa

Gorrotu dut Kepak jakin gabe hitzegite hori

kepari buruz ezer jakin gabe etengabe hizketan aritza gorrotatu egiten dut [g]

gorrota egiten dot kepen jakin barik berba egiteko modua/kepak jakin barik berba egitea [b]

40. Quiero ese tranquilo hablar de Kepa

\*Kepak lasai hitzegite hori nahi dut

keparen hizketa lasai hori maite dut [g]

kepen berba egiteko modu trankil hori nahi dot [b]

305

\*Gobernuak eta etak suetena noiz hasi eta presoak noiz askatu  
negoziatuko dituzte  
Gobernuak eta etak presoak askatzea eta armak betirako uztea  
negoziatuko dituzte  
Gobernuak eta etak presoen askapena eta indarkeriarekin emaiera  
negoziatuko dute  
Gobernuak eta e.t.a.k presoen askatasuna eta biolentziaren amaiera  
negoziatuko dute [g]  
gobernuak eta ETAk presuen libraketea eta biolentzien amitukerea  
negoziatuko dabe [b]

49. El gobierno y eta negociarán cuándo empezar el alto el fuego y  
cuándo liberar a los presos  
gobernuak eta e.t.a.k noiz hasi su etena eta noiz askatu presoak  
negoziatuko dute [g]  
gobernuak eta ETAk "alto el fuego" a noiz hasi eta presoak noiz libraturu  
negoziatuko dabe b]

50. Hemos descubierto que ser intelibente y trabajar duro no es  
suficiente  
azkarra izatea eta gogor lan egitea ez dela nahikoa aurkitu dugu [g]  
inteligentea izatea eta biarra gogor egitea ez dela nahiko ikasi dogu [b]

51. Me has vendido el libro  
Zuk niri liburua saldu didazu  
liburua saldu didazu [g]  
liburua saldu deustazu [b]

52. Me has vendido al carnicero  
\*Zuk harakinari ni saldu n(a)iozu  
harakinari saldu nauzu [g]  
karnizeruari saldu nauzu [b]

53. Me parece mal que me vendas al carnicero  
Gaizki iruditzen zait zuk ni harakinari saltzea

harakinari zuk ni saltza gaizki iruditzen zait [g]  
txarto ereitzen jat karnizeruari ni saltza [b]  
54. Espero la venida de Ainhoa a Seattle  
Espero dut Ainhoa Seattlera etortzea  
Espero dut Ainhoa Seattlera etorria  
ainhoaaren ettorera seattlerea espero dut (ainhoa seattlera noiz etorriko  
den zai nago[jatorrago]) [g]  
ainhoa seattlera etortea itxoten nago [b]  
55. Espero que Ainhoa haya venido a Seattle  
ainhoa seattlera etorrita egotea espero dut [g]  
ainhoa seattlera etorri dala esperantzia daukat [b]  
56. ¿Quieres ayuda?  
Nahi duzu laguntza  
laguntza nahi al duzu? [g]  
laguntza nahi duzu? /laguntzarik nahi dozu? [b]  
57. A Mikel se le olvidó traer el libro  
Mikelek ahaztu du liburua ekartzea  
mikeli liburua ekartzea ahaztu egin zitzaion [g]  
mikeleri liburua ekartea aztu jakon [b]  
58. A Mikel se le olvidó que trajeras el libro  
\*Mikelek ahaztu du zuk liburua ekartzea  
mikeli ahaztu egin zitzaion zuk liburua ekartzea [g]  
mikeleri zuk liburua ekartzea aztu jakon [b]  
59. Su escritura (el hecho de que escriba)  
Hark idaztea  
bere idazkera [g]  
berak idaztea [b]  
60. Su escritura (su documento)  
Haren idaztea  
bere idatzia [g]

## APPENDIX B

### ENGLISH SENTENCES GIVEN TO SUBJECTS

1. It is bad to have killed Calvo in prison
2. I regret that father is ill
3. I regret father's being ill
4. It is easy to read this book
5. Father lets me speak
6. Mother hasn't succeeded in getting the child to eat
7. I forgot how to write
8. As soon as Gorka entered the room, we all left
9. She intended to have dinner
10. Father's being so well surprised me
11. Mother's purchases have been good
12. The purchase of the books and the dresses
13. Father's purchase of the car was timely
14. Grand father's calm walking pleased me
15. Grand father's working silently surprised me
16. The type of driving that grand father liked was dangerous
17. I don't like your comings and goings
18. Your night walks have to stop
19. The arrival of the man
20. The weather in England
21. England's weather
22. Father's arrival of yesterday
23. I decided to cut my beard
24. The solving of problems by the students
25. I decided to be silent
26. We know how to bother people
27. I didn't dare to say anything
28. He said that he saw a cow flying
29. They forbid me to speak at meetings
30. Your going to America made me happy
31. Your going to America doesn't make me happy
32. The fact that Miren has arrived late is surprising
33. I hate the fact that Miren cannot be quiet
34. The fact that the children move continuously makes me nervous
35. I want Kepa to talk calmly
36. Ainhoa worries me terribly
37. What to do next year worries me
38. Ainhoa's arrival worries me very much
39. It is necessary for Ainhoa to bring beer
40. In whose opinion is it necessary for Ainhoa to bring beer
41. In whose opinion is Ainhoa's bringing beer necessary?
42. The president demanded the liberation of prisoners and the negotiation of a cease-fire
43. The government and ETA will negotiate the liberation of the prisoners and the end of violence
44. We have discovered that to be smart and to work late nights and weekends is not enough
45. You have sold me the book
46. You have sold me to the butcher
47. It seems wrong for you to sell me to the butcher
48. I expect Ainhoa to come to Seattle
49. Do you want (me) to help you?

Anderson, S.R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Andrews, A. 1985. The major functions of the noun phrase. In: Shopen (ed.), vol.1: 62-124.

Appel, R. and Muysken, P. 1987. *Language contact and bilingualism*. London, Baltimore: Edward Arnold.

Aronoff, M. 1976. *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press.

Arrieta, K., Joseph, B.D., Smirniotopoulos, J. 1986. How Ergative is Basque? Eastern States Conference on Linguistics, ed. by Marshall, F., Miller, A. & Zheng-sheng Zhang. 25-37. Pittsburgh: University of Pennsylvania.

Artiagoitia, Xabier. 1992. *Verbal Projections in Basque and Minimal Structure*. Ph.D. Dissertation. University of Washington. Seattle, Washington.

Artiagoitia, Xabier. 1996. Oral Presentation. University of Deusto. Deusto, Basque Country, Spain.

Azkarate, M. 1993. *Basque Compound Nouns and Generative Morphology: Some Data*. In: Hualde, J.I. & Ortiz de Urbina, J. (eds.), p. 221-243.

Azkue, R.M. 1923. *Morfología Vasca*. Bilbao, Vizcaya.

Bach E. 1968 . The noun phrase. *Universals of Linguistic Theory*, ed by E. Bach & R.T. Harms. New York: Holt, Rinehart, and Winston.

Baetens Beardsmore, H. 1982. *Bilingualism: Basic principles*. Clevedon.

Baker M. 1985. *Syntactic Affixation and English Gerunds*. Proceedings of the West Coast Conference of Formal Linguistics, ed by Cobler et al. Palo Alto, CA: Stanford University.

Bartning, I. 1986. Aspects des syntagmes binominaux en de en français. *Travaux de linguistique et de littérature* xxiv, 1. 347-71.

Berko-Gleason, J. 1982. *Insights from Child Language Acquisition for Second Language Loss. The Loss of Language Skills*, ed. by Lambert, R.D. & B. F. Freed.

Bierwisch, M. 1989. *Event Nominalization: Problems and Proposals. Wortstruktur und Satzstruktur*, ed. by W. Motsch. Akademie der Wissenschaften der DDR.

Bloomfield, L. 1933. *Language*. New York: Holt.

Blount, Ben G. 1974. *Language, Culture and Society: A Book of Readings*. Cambridge, Mass.: Winthrop.

Blount, B.G. & M. Sanches. 1977. *Introduction: Sociocultural Dimensions of Language Change*. *Sociocultural Dimensions of Language Change*, ed. by Blount, B.G. & M. Sanches. New York: Academic Press.

Bolkenstein, A.M., de Groot, C., and Mackenzie, J.L. (eds.) 1985. *Predicates and Terms in Functional Grammar*. Dordrecht, Holland: Foris Publications.

Borer, Hagit. 1993. *Parallel Morphology*. Book Manuscript. University of Massachusetts at Amherst.

Bošković, Željko. 1996. Selection and the Categorial Status of Infinitival Complements. *NLLT* 14. 269-304.

Bowers J.S. 1975. Some Adjectival Nominalizations in English. *Lingua* 37. 341-361.

Bresnan, Joan . 1982. *Control and Complementation. The Mental Representation of Grammatical Relations*, ed. by Bresnan, J. Cambridge, MA: The MIT Press.

Bresnan, Joan. 1995. *Lexicality and Argument Structure*. Paris Syntax and Semantics Conference, October 12. Paris.

Bresnan, Joan. 1997. *Mixed Categories as Head Sharing Constructions*. Proceedings of the LFG97 Conference. University of California, San Diego.

Breton, R.J.L. 1991. *Geolinguistics. Language Dynamics and Ethnolinguistic Geography*. Ottawa: University of Ottawa Press.

Comrie, B. 1978. Ergativity. *Syntactic Typology*, ed. by Lehmann. 329-94.

Comrie, B. 1981. *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell.

Comrie, B. and Thompson, S.A. 1985. Lexical Nominalization. In: Shopen (ed.).

Corbett, G.G. 1987. The Morphology/Syntax Interface: Evidence from Possessive Adjectives in Slavonic. *Language* 63, 2. 299-345.

Dahl, O. 1980. Russian Direct Objects and Functional Case Marking Principles. *Slavica Gothoburgensis* 7, ed. by G. Jacobson. 17-27. Gothenburg: Gothenburg University.

Davidson D. 1980. *Essays on Actions and Events*. Oxford: Clarendon Press.

De Miguel, Elena. 1995. An Aspectual Restriction on Spanish Nominal Infinitives. *ASJU International Journal of Basque Linguistics and Philology* XXIX-1. Donostia.

De Miguel, Elena. 1996. Nominal Infinitives in Spanish: An Aspectual Constraint. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 41, 1. 29-53

Dench, A. and Evans, N. 1988. Multiple Case-marking in Australian Languages. *Australian Journal of Linguistics* 8. 1-47.

Denison, N. 1977. Language Death or Language Suicide. *Linguistics* 191. 13-22.

Denison, N. 1982. A linguistic ecology for Europe? *Folia Linguistica* 16. 5-16.

Denison, N. 1989. Twist the Scylla of total assimilation and the Charybdis of suicidal purism. *York Papers in Linguistics* 13. 101-113.

Dik, S.C. 1985. Formal and Semantic Adjustment of Derived Constructions. In: Bolkenstein, A. M., et al. (eds.).

Dixon, R. M. W. (ed.) 1976. *Grammatical categories in Australian languages*. Canberra : Australian Institute of Aboriginal Studies ; [Atlantic Highlands] N.J. : Humanities Press.

Dixon, R.M.W. 1979. Ergativity. *Language* 55. 59-138.

Dorian, Nancy. 1973. Grammatical change in a dying dialect. *Language* 49. 413-438.

Dorian, Nancy. 1977. The Problem of the Semi-Speaker in Language Death. *Linguistics* 191. 23-32.

Dorian, Nancy. 1978. The Fate of Morphological Complexity in Language Death: Evidence from East Sutherland Gaelic. *Language* 54, 3. 590-609.

Dorian, Nancy. 1981. *Language Death: The Life Cycle of a Scottish Gaelic Dialect*. Philadelphia: University of Pennsylvania Press.

Dorian, Nancy. 1982a. *Language Loss and Maintenance in Language Contact Situations. The loss of language skills*, ed. by R.D. Lambert, B.F. Freed. New York: Rowley.

Dorian, Nancy. 1982b. *Linguistic models and language death evidence. Exceptional Language and Linguistics*, ed. by L.K. Obler, L. Menn. New York: Academic Press.

Dorian, Nancy. 1983. Natural second language acquisition from the perspective of the study of Language Death. *Pidginization and Creolization as Language Acquisition*, ed. by Roger W. Andersen. Mas.: Rowley.

Dorian, Nancy. 1985. Vocative and imperative in decline. *Studia Linguistica Diachronica et Synchronica*, ed. by Pieper & Stckel. 161-174. New York: Mouton de Gruyter.

Dorian, Nancy. 1986. Making do with less: some surprises along the language death proficiency continuum. *Applied Psycholinguistics* 7. 257-276.

Dorian, Nancy (ed.). 1989. *Investigating Obsolescence. Studies on Language Contraction and Death*. Cambridge: Cambridge University Press.

Fillmore, C.J. 1971. Some Problems for Case Grammar. Report of the Twenty-Second Annual Round Table Meeting on Linguistics and Language Studies, ed. by R. J. O'Brien. 35-56. Washington, D.C.

Fishman, J. A. 1964. Language Maintenance and Language Shift as a Field of Inquiry. *Linguistics* 9. 32-70.

Fishman, J. A. 1967. Bilingualism With and Without Diglossia; Diglossia With and Without Bilingualism. *Journal of Social Issues* XXIII, 2. 29-38.

Fishman, J. A. 1972a. The sociology of language: an interdisciplinary social science approach to language in society. Newbury House Publishers.

Fishman, J. A. 1972b. Advances in the Sociology of Language II. The Hague: Mouton.

Fishman, J. A. 1977. Language and Ethnicity. In: Giles, H. (ed.), *Language, Ethnicity and Intergroup Relations*. New York: Academic Press.

Fishman, J. A. 1989a. Bilingualism and Biculturalism as Individual and as Societal Phenomena. *Language and Ethnicity in Minority Sociolinguistics Perspective*, ed. by Fishman, J. 181-202. Philadelphia: Multilingual Matters Ltd., Clevedon.

Fishman, J. A. 1989b. Language Maintenance and Ethnicity. *Language and Ethnicity in Minority Sociolinguistics Perspective*, ed. by Fishman, J. 202-224. Philadelphia: Multilingual Matters Ltd., Clevedon.

Fishman, J. A. 1989c. The Societal Basis of the Intergenerational Continuity of Additional Languages. *Language and Ethnicity in Minority Sociolinguistics Perspective*, ed. by Fishman, J. 224-233. Philadelphia: Multilingual Matters Ltd., Clevedon.

Fishman, J. A. 1989d. The Spread of English as a New Perspective for the Study of Language Maintenance and Language Shift. *Language and Ethnicity in Minority Sociolinguistics Perspective*, ed. by Fishman, J. 233-265. Philadelphia: Multilingual Matters Ltd., Clevedon.

Fishman, J. A. 1991. Reversing Language Shift. Theoretical and Empirical Foundations of Assistance to Threatened Languages. Philadelphia: Multilingual Matters. Ltd., Clevedon.

Foley, W. 1988. Language birth: the processes of pidginization and creolization. In: Newmeyer.

Foley, W.A. and Van Valin, R.D., Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Frazier, Lyn. 1985. *Syntactic Complexity. Natural Language Parsing*, ed. by Karttunen, L & A. Zwicky.

Frazier, Lyn. and Janet Fodor. 1978. The Sausage Machine: A New Two-Stage Parsing Model. *Cognition: International Journal of Cognitive Science* 6. 291-325.

Frazier, Lyn, Charles Clifton and Janet Randall. 1983. *Filling Gaps: Decision Principles and Structure in Sentence Comprehension*. *Cognition: International Journal of Cognitive Science* 13, 2. 187-222.

Freed, B.F. 1982. Language Loss: Current thoughts and future directions. *The Loss of Language Skills*, ed. by Lambert, R.D. & Freed, B. A.

Fu, Jingqi. 1994. On Deriving Chinese Derived Nominals: Evidence for V-to-N Raising. Ph.D. Dissertation. University of Massachusetts. Amherst.

Fullana, O. 1995. Why Nominal Infinitives Express Manner. *Catalan Working Papers in Linguistics (CatWPL)* 4.2. 221-227.

Gaarder, B.A. 1979. Language maintenance or language shift. *Bilingualism in early childhood*, ed. by Mackey, W.F. and Andersson, T. 409-34. Mas.: Rowley.

Gal, Susan. 1979. Language shift: Social determinants of linguistic change in bilingual Austria. New York.

Gal, Susan. 1988. The Political Economy of Code Choice. *Codeswitching: Anthropological and Sociolinguistic Perspectives*, ed. By Monica Heller. 245-264.

Gardner, Nick. 1998. Basque in Education in the Basque Autonomous Community. ms.

Hopper, P.J. and Thompson, S.A. 1980. Transitivity in Grammar and Discourse. *Language* 56. 251-99.

Horn G. 1975. On the Nonsentential Nature of the Poss-Ing Construction. *Linguistic Analysis* 1, 4.

Hualde, J.I. 1989. A Lexical Phonology of Basque. Dissertation Abstracts International. Ann Arbor, Michigan.

Hualde, J.I. and J. Ortiz de Urbina (eds.) 1993. Generative Studies in Basque Linguistics. Amsterdam: John Benjamin's Publishing Co.

Hymes, D. 1964. Language in Culture and Society: A Reader in Linguistics and Anthropology. New York: Harper & Row.

Hymes, D. 1967. Models of the Interaction of Language and Social Setting. *Journal of Social Issues* XXIII, 2. 8-28.

Hymes, D. 1974a. The Ethnography of Speaking. Language, culture and society, ed. by Blount, B. G.

Hymes, D. 1974b. Sociolinguistics and the Ethnography of Speaking. In: Blount, B. G. (ed.)

Hymes, D. 1983. Essays in the History of Linguistic Anthropology. Amsterdam: John Benjamins Publishing Co.

Jackendoff, R. 1977. X' Syntax: A Study of Phrase Structure. Cambridge: MIT Press.

Jeffers, R.J. 1976. Typological Shift and Change in Complex Sentence Structure. In: Steever, S.B. et al. (eds.). 136-49.

Johns, A. 1992. Deriving ergativity. *Linguistic Inquiry* 23. 57-87.

Johnson, K. 1988. Clausal Gerunds, the ECP, and Government. *Linguistic Inquiry* 19, 4. 583-611.

Jokinen, K. 1991. On the Two Genitives in Finnish. In: EuroTyp Working Papers (Programme in Language Typology, European Science Foundation), Series 7, no. 14. Strasbourg: European Science Foundation.

Joppen, S & D. Wunderlich. 1995. Argument Linking in Basque. *Lingua* 97. 123-171.

Joseph, B. D. 1996. Diachronic Perspectives on Control. Control and Grammar, ed. by Larson, R.D., Iatridou, S., Lahiri, U. and Higginbotham, J. 195-234. Dordrecht: Kluwer Academic Publishers.

Joseph, B. D. 1998. On some control structures in Hellenistic Greek. ms.

Kamareddine, F. and E. Klein. 1992. Nominalization, Predication and Type Containment. *Journal of Logic, Language and Information*.

Kamp, H. and U. Reyle. 1994. From Discourse to Logic. Dordrecht: Kluwer.

Kapeliuk, O. 1980. Evolution de la phrase amharique: la nominalization du verbe. Modern Ethiopia from the Accession of Menilek II to the Present, ed. by Tubiana, J. 97-106. Rotterdam: A. A. Balkema.

Kaplan, Ronald M and Joan Bresnan. 1982. Lexical Functional Grammar: A Formal System for Grammatical Representation. The Mental Representation of Grammatical Relations, ed. by Bresnan, Joan. Cambridge: The MIT Press.

Karttunen, F. 1977. Finnish in America: A Case Study in Monogenerational Language Change. Sociocultural Dimensions of Language Change, ed. by Blount, B.G. & M. Sanches. New York: Academic Press.

Karttunen, L. 1986. Radical Lexicalism. Report No. CSLI-86-68. CSLI : Stanford.

Kayne R. 1984. Unambiguous Paths. Chapter 7 of Connectedness and Binary Branching. Dordrecht: Foris Publications.

Keenan, E.L. 1974. Generalizing the Notion of 'Subject of. Chicago Linguistic Society 10. 298-309.

Keenan, E.L. 1976. Towards a Universal Definition of Subject. In: Li (ed.). 303-33.

López Morales, H. 1993. *Sociolinguística*. Madrid: Gredos.

Lord, C. 1976. Evidence for Syntactic Reanalysis: from Verb to Complementizer in Kwa. In: Steever et al. (eds.). 179-91.

Lyons, Ch. 1986. The Syntax of English Genitive Constructions. In: *Journal of Linguistics* 22. 123-43.

Maandi, Katrin. 1989. Estonian among Immigrants in Sweden. *Investigating Obsolescence: Studies in Language Contact and Death*, ed. by Nancy C. Dorian. 227-241.

Mackenzie, J.L. 1985. Nominalization and Valency Reduction. In: Bolkenstein et al. (eds.). 29-47.

Mackey, W.F. 1983. U.S. Language Status Policy and the Canadian Experience. *Progress in Language Planning*, ed. by Cobarrubias, J & J. Fishman.

Mackey, W. and Ornstein, J. (eds.) 1979. *Sociolinguistic studies in language contact*. The Hague.

Makino S. 1968. *Some Aspects of Japanese Nominalizations*. Tokyo: Tokai University Press.

Marantz, A. 1978. Embedded sentences are not noun phrases. *Proceedings of the Eighth Annual Meeting of the North Eastern Linguistic Society*, ed. by M.J. Stein. 112-122. Amherst: Graduate Linguistic Student Association, University of Massachusetts.

Martin-Jones, M & S. Romaine. 1986. Semilingualism: a half-baked theory of communicative competence. *Applied Linguistics* 7, 1.

Martinet, A. 1982. Bilinguisme et diglossie. *La Linguistique* 18. 5-16.

Meisel, J. 1980. Linguistic simplification. Second language development: Trends and issues, ed. by Felix, S.W. 13-40. Tübingen.

Matisoff, J.A. 1972. Lahu Nominalization, Relativization and Genitivization. *Syntax and Semantics*, vol. 1, ed. by Kimball, J.P. 237-58. New York: Seminar Press Inc.

Morales, A. 1979. Estructuras sintácticas anglicadas en el español de Puerto Rico: infinitivos y gerundios (análisis transformacional). *Boletín de la Academia Puertorriqueña de la Lengua Española* 7, 2. 111-128.

Mugica, L.M. 1978. *Origen y desarrollo de la sufijación euskérica*. San Sebastián: Ediciones Vascas.

Mühlhäuser, P. 1974. *Pidginization and simplification of language*. Canberra.

Munro, P. 1982. On the Transitivity of "Say" Verbs. *Studies in Transitivity*, ed. by Hopper, P. J. and Thompson, S.A. 301-18. New York: Academic Press.

Muysken, P. 1981. Halfway between Quechua and Spanish: the case of relexification. *Historicity and variation in creole studies*, ed. by Highfield, A. and Valdman, A. 52-78. Ann Arbor, Michigan.

N'Diaye, G. 1970. *Structure du dialecte basque de Maya*. The Hague: Mouton.

Newmeyer, F.J. 1970a. The Source of Derived Nominals in English. *Language* 47. 786-796.

Newmeyer, F.J. 1970b. The Derivation of the English Action Nominalization. *Papers from the 6th Regional Meeting of the Chicago Linguistic Society*. 408-415.

Newmeyer, F.J. (ed.) 1988. *Linguistics: The Cambridge Survey*. Vol. IV *Language: The Socio-cultural context*. Cambridge: Cambridge University Press.

Nichols, J. 1986. Head-marking and Dependent-marking Grammar. *Language* 62, 1. 59-119.

Nichols, J. 1988. Nominalization and Assertion in Scientific Russian Prose. In: Haiman and Thompson (eds.). 399-428.

Obler, L.K. 1982. *Neurolinguistic Aspects of Language Loss as they Pertain to Second Language Attrition. The Loss of Language Skills*, ed. by Lambert, R.D & B.F. Freed.

Rosen, S.T. 1989. Two Types of Noun Incorporation: A Lexical Analysis. *Language* 65, 2. 294-317.

Ross J. 1973. Nouniness. Three Dimensions of Linguistic Theory, ed. by O. Fujimura. Tokyo: TEC Company.

Saltarelli, M. 1988. Basque. Croom Helm Descriptive Grammars Series. New York: Croom Helm.

Sankoff, D. 1988. Sociolinguistics and syntactic variation. In: Newmeyer (ed.).

Scalise, S. 1984. Generative Morphology. Dordrecht: Foris.

Schachter, P. 1976. A nontransformational analysis of gerundive nominals in English. *Linguistic Inquiry* 7. 205-241.

Schmidt, A. 1985. The fate of ergativity in dying Dyirbal. *Language* 61. 378-396.

Schwartz, A. 1970. On Interpreting Nominalizations. *Progress in Linguistics*, ed. by Bierwiesch, M and Heidolph, K.E. 295-305.

Selkirk E. 1977. Some Remarks on Noun Phrase Structure. *Formal Syntax*, ed. by P. Culicover, T. Wasow, & A. Akmajian. New York: Academic Press.

Selkirk, E.O. 1982. The Syntax of Words. Cambridge: The MIT Press.

Sells, Peter. 1995. Korean and Japanese morphology from a lexical perspective. *Linguistic Inquiry* 26. 277-325.

Shibatani, M. 1977. Grammatical Relations and Surface Cases. *Language* 53, 4. 789-810.

Shopen, T. (ed.) 1985. *Language Typology and Syntactic Description*, vol. III. Cambridge: Cambridge University Press.

Silva Corvalán, Carmen. 1994. Language Contact and Change. Spanish in Los Angeles. Oxford: Clarendon Press.

Silverstein, M. 1976. Hierarchy of Features and Ergativity. In: Dixon (ed.). 112-171.

Slobin, Dan I. 1977. *Language Change in Childhood and in History*. *Language Learning and Thought*, ed. by John Theodore Macnamara. 185-214.

Steele, S. 1978. Word Order Variation: A Typological Study. In: Greenberg et al. (eds.), vol4. 585-623.

Spencer, Andrew. 1992. *Morphological Theory*. Oxford: Blackwell.

Spencer, Andrew. 1995. On the Typology of Deverbal Nominalizations. *Working Papers in Language Processing* 48. Department of Language and Linguistics. University of Essex.

Stati, S. 1989. La nominalisation des infinitifs. In: *Studi Italiani di Linguistica Teorica ed Applicata*.

Subramanian, U. 1992. On Internal and External Syntax Mismatches. Ph.D. Dissertation. The Ohio State University.

Swadesh, Morris. 1948. Sociologic notes on obsolescent languages. In: *International Journal of American Linguistics* XIV. 26-235.

Szabolcsi, A. 1987. Functional Categories in the Noun Phrase. *Approaches to Hungarian*, vol. 2, ed. By I. Kenesei. Szeged.

Tabouret-Keller, A. 1968. Sociological factors of language maintenance and language shift: A methodological approach based on European and African examples. *Language Problems of Developing Nations*, ed. by J.A. Fishman et al. 107-118. New York: Wiley.

Thomason, Sarah G. 1986. On Establishing External Causes of Language Change: State Univ. of New York at Buffalo, Buffalo, New York, Oct 3-5, 1985. *Proceedings of the Second Eastern Conference on Linguistics*, ed. by Choi, Sonia, Dan Devitt, Wynn Janis, Terry McCoy and Zheng-sheng Zhang. 243-251.

Thomason, Sarah Grey & Terrence Kaufman. 1991. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Thompson S. 1973. On Subjectless Gerunds in English. *Foundations of Language* 9. 374-383.

Zucchi, A. 1993. *The language of Propositions and Events. Issues in the Syntax and the Semantics of Nominalization*. Dordrecht: Kluwer Academic Publishers.

Yoon, James Hye Suk. 1996. Nominal gerund phrases in English as phrasal zero derivations. *Linguistics* 34. 329-356.

Yoon, J. H. & N. Bonet-Farran. 1991. The Ambivalent Nature of Spanish Infinitives. *New Analyses in Romance Linguistics*, ed. by Wanner D. & D.A. Kibbee. Amsterdam: John Benjamins Pu. Co.

Zwicky, A. 1985. Heads. *Journal of Linguistics* 21. 1-29.

Zwicky, A. 1992. Jottings on Adpositions, Case Inflections, Government and Agreement. *The Joy of Grammar: A Festschrift for James D. McCawley*, ed. by D. Brentari, G.N.Larson and L.A. MacLeod . Chicago: University of Chicago Press.