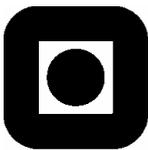

Second Language Acquisition and Processing of Norwegian DP Internal Agreement

Fufen Jin

Doctoral Thesis

Submitted for the Partial Fulfilment of the Requirements for the Degree of

Philosophiae Doctor



Department of Scandinavian Studies and Comparative Literature
Faculty of Arts
Norwegian University of Science and Technology

March 2007

Copyright © 2007 Fufen Jin

ISBN printed 978-82-471-1514-5

ISBN electronic 978-82-471-1528-2

ISSN 1503-8181

NTNU 2007: 69

Printed in Norway by NTNU Trykk, Trondheim

TO MY PARENTS FENGSCHAN JIN AND XUEXIANG LI,
MY HUSBAND JINGYUE LI AND MY SON RUIYI LI

Abstract

The primary goal of this study is to determine whether adult second language learners ultimately achieve linguistic competence that is identical to that of native speakers, not only at the level of grammatical knowledge but also at the level of processing. This issue is addressed through investigating the production and perception of Norwegian Determiner Phrase (DP) internal agreement by endstate L2 speakers of Norwegian. Participants in the experiment include three different L2 groups (L1 English, L1 Italian/Spanish, and L1 Chinese), and a native control group.

The target language, Norwegian, has DP-internal agreement between determiners, attributive adjectives, and nouns in terms of gender, number, and definiteness. Based on Julien's (2005) DP model, D and α heads in Norwegian are assumed to carry uninterpretable gender, number and definiteness features. The source languages in question differ parametrically from the target language with respect to one, two or all of the uninterpretable features relevant for DP internal agreement. The informants are tested on two tasks: an online DP production task, and an auditory naming task. The purpose of the DP production task is to determine whether the uninterpretable features will be present in the interlanguage grammar regardless of status of the corresponding features in the L1s; the purpose of the auditory naming task is to examine whether L2 learners will be sensitive to the agreement cues on D when processing L2 Norwegian, and whether their sensitivity (or insensitivity) to the agreement cues is directly related to how well they produce the correct agreements.

Results from the on-line DP production task divide the L2 participants into two groups: some informants (referred to as G1 participants) performed in a nativelike way, while the others (referred to as G2 participants) showed a strong L1 effect—those uninterpretable features which are not instantiated in L1s are areas of persistent difficulty in these L2 speakers. Results from the auditory naming task show that endstate L2 speakers do not process Norwegian DP internal agreement in a nativelike fashion. While native speakers are sensitive to the agreement cues on D, L2 speakers are totally insensitive to the grammatical agreement violations in L2 grammar. This observation indicates that L2 processing is less automatic and involves less full syntactic computation than L1 processing. A comparison of G1 and G2 participants reveals a remarkably similar processing pattern — neither group shows sensitivity to the agreement cues on D, contra to what has been found in native speakers. This finding indicates that adult L2 speakers have not attained nativelike processing, even after they have attained nativelike grammatical knowledge, suggesting that certain processing mechanisms are subject to a critical period.

Preface

This thesis is submitted to the Norwegian University of Science and Technology (NTNU) for partial fulfilment of the requirements for the degree of Philosophiae Doctor.

The work referred to has been performed at the Department of Scandinavian Studies and Comparative Literature, NTNU, Trondheim, under the supervision of Professor Tor A. Åfarli.

Acknowledgements

No words can express my gratitude to my supervisor, Professor Tor A. Åfarli, who is virtually responsible for every step of my intellectual growth during my three-year Ph.D program. His vast knowledge of syntax and second language acquisition has inspired me a lot. A discussion with him is more often worth ten years of reading by myself. He read the manuscript of the thesis word by word. His careful and insightful comments helped me interpret syntactic theories correctly, present the thesis more clearly, and remove many grammatical errors.

With deepest gratitude, I thank Dr. Kristin E. Eide for her friendship, encouragement, and continuous support. I always admire her research competence and positive, cheerful personality. With her expertise on generative syntax, she has provided invaluable input to my studies. With her sunny smiles and warm friendship, she has helped me keep in good spirits even in cold, dark winter days. Her dedicated work in organizing the Språkmøter seminar and the Minimalist Syntax seminar has created a stimulating syntactic research environment at NTNU, which benefit linguistic researchers and students alike, including me.

I gratefully acknowledge Dr. Chris Wilder for allowing me to sit in his syntax course and for numerous discussions about the DP syntax in the early stage of my thesis. Special thanks are due to Torbjørn Nordgård, Jørn Almberg and Brit Mæhlum for teaching me the methodology course.

I am grateful to our Språkmøter project research team for their cooperation and support. Our project manager, Professor Wim van Dommelen, deserves my special thanks for his technical support during the experiments and for his help with statistic analysis. I also wish to express my gratitude to the informants who participated in the experiments and patiently answered my questions, to Randi A. Nilsen for her help with the recording, to Guro Busterud and Marthe-Julie Lundeby for their help in interviewing the L2 informants, and to Rein Ove Sikveland for his help in running the experiment and checking the transcriptions.

Many researchers outside NTNU have also helped me a lot with their insights, advice, or course supervision. I am most grateful to Professor Roger Hawkins and Dr. Ingrid Leung for their kindness and hospitality during my research visit to the University of Essex in spring, 2005. I also feel indebted to members of the second language acquisition discussion group and to my course supervisors, Ingrid Leung, Claudia Felser, Bob Borsley, Martin Atkinson and Andrew Radford. Thanks are also due to Marit Julien at Lund University for answering my questions regarding the DP structure of Norwegian. Dr. Boping Yuan at Cambridge University visited NTNU in 2006 and gave a series of lectures on second language acquisition, from which I gleaned much needed inspiration for my own study. I also thank him for his very encouraging comments on my thesis. Different parts of this work (in collaboration with Tor A. Åfarli and Wim van Dommelen) had been presented at Eurosla 16 and GALANA-2. I thank the audience for their questions and insightful comments.

I wish to thank my Norwegian language teachers Sissel Robins, Åsta Øvregaard for their instruction, and Sissel Furuseth for providing me the material in designing the Norwegian proficiency test. Special thanks go to Olaf Husby for very pleasant cooperation in writing the co-authored book, *Innføring i kinesisk for nordmenn og innføring i norsk for kinesere*.

I would like to express my gratitude to the administrative personnel in the HF faculty and departmental secretaries for creating a pleasant work environment. I thank my colleagues and friends, especially Ragnhild Ljosland, Annika Odland, Gunnhild Aamodt, and Mari Nygård, for being so kind to me. I also thank my Chinese friends in Trondheim, especially Naiquan Ye, Shanshan Jiang, Hua Liu, Yanjun Li, Jingdong Liu, Shaoli Yang, and Heng Yang, for their care and friendship.

My final word of gratitude is reserved for my beloved husband Dr. Jingyue Li and my son Ruiyi Li. A thousand thanks to them for making home a paradise.

Contents

Abstract.....	i
Preface.....	iii
Acknowledgements	v
Contents	vii
List of Figures.....	x
List of Tables	x
Abbreviation list.....	xiii
1 Overview.....	1
1.1 Introduction	1
1.2 Research questions	4
1.3 Thesis outline	5
2 Morphosyntax of DPs in the target and the source languages	7
2.1 Introduction	7
2.2 A comparison of DPs of the target and source languages	7
2.3 Gender and number systems of Norwegian and Italian/Spanish	12
2.3.1 Gender and number in Norwegian.....	12
2.3.2 Gender and number in Italian/Spanish	18
2.4 Summary	21
3 Theoretical assumptions.....	23
3.1 Introduction	23
3.2 Some major claims of the MP	23
3.3 Syntactic structure of Norwegian DPs	27
3.4 Syntactic structure of Italian/Spanish DPs	35
3.5 Syntactic structure of English DPs	37
3.6 A note on Chinese DPs	38
3.7 Summary and a road ahead	39
4 Literature review	41
4.1 Introduction	41
4.2 Failed Functional Features Hypothesis (FFFH)	42
4.2.1 Hawkins and Chan (1997)	42

4.2.2	Hawkins (2004); Hawkins and Hattori (2006)	45
4.2.3	Franceschina (2001)	48
4.3	<i>Full Transfer Full Access (FTFA)</i>	49
4.3.1	Schwartz and Sprouse (1994, 1996)	49
4.3.2	White <i>et al.</i> (2004)	51
4.3.3	Goad <i>et al.</i> (2003), Goad and White (2006)	54
4.4	<i>Shallow Structure Hypothesis (SSH)</i>	58
4.5	<i>A summary</i>	62
5	L2 production.....	65
5.1	<i>Introduction</i>	65
5.2	<i>The participants</i>	66
5.2.1	The L2 participants	66
5.2.2	Norwegian Proficiency Test (NPT)	69
5.2.3	L1 participants	70
5.3	<i>Experiment 1: Gender Assignment</i>	71
5.3.1	Introduction	71
5.3.2	Method	71
5.3.3	Results	72
5.4	<i>Experiment two: On-line DP production task</i>	74
5.4.1	Introduction	74
5.4.2	Predictions	74
5.4.3	Method	77
5.5	<i>Results</i>	80
5.5.1	Rate of missed items	80
5.5.2	L1 group	80
5.5.3	L2 groups	80
5.6	<i>Summary of findings and discussion</i>	91
5.6.1	Summary of findings from experiment one and two	91
5.6.2	Predictions revisited	92
5.6.3	Accounting for L1-L2 differences in production	95
6	L1 & L2 Perception	99
6.1	<i>Introduction</i>	99
6.2	<i>L1 and L2 feature processing</i>	100
6.3	<i>The experiment</i>	105
6.3.1	Method	106
6.3.2	Results	110
6.4	<i>Summary of findings and discussion</i>	117
6.4.1	A summary of L1 & L2 processing results	117
6.4.2	The locus of priming effects in L1 processing	118
6.4.3	Agreement processing mechanisms	119
6.4.4	L1/L2 processing differences	121
6.4.5	Is there an L1 transfer effects on L2 processing?	124
6.5	<i>Conclusion</i>	125
7	Conclusions and future directions.....	127

7.1	<i>Summary and conclusion</i>	127
7.2	<i>Contributions</i>	128
7.3	<i>Directions for future work</i>	130
References.....		133
Appendix.....		143

List of Figures

Figure 4.1. The FFFH's and the FTFA's accounts for L2 ultimate attainment.....	62
Figure 4.2. L2 processing model in accordance with the SSH.....	63
Figure 5.1. Each L2 group's accuracy on the gender assignment task.....	74
Figure 5.2. Accuracy rate of the DPs produced by the each L2 informant	82
Figure 5.3. Accuracy on D and <i>a</i> for each L2 participant	82
Figure 5.4. Individual L2 group' error rate (%) on each of the categories.....	85
Figure 6.1. The L1 group's mean RTs to concordant and discordant neuter and common gender-marking.....	113
Figure 6.2. The L2 participants' mean RTs to concordant and discordant neuter and common gender-marking.....	115

List of Tables

Table 2.1. Inflectional categories and agreement in DPs of the L2 and the L1s.....	12
Table 2.2. Inflection paradigm of modified Norwegian DPs	18
Table 2.3. Illustrations of DP internal agreement in the L2 and L1s	21
Table 3.1. Parameterized uFs in the target and the source languages	40
Table 4.1. Chinese group's accuracy in judging English RRCS (%).....	43
Table 4.2. Chinese group's accuracy in rejecting <i>wh</i> -island and complex NP violations	43
Table 4.3. French group's accuracy in judging English RRCS (%).....	44
Table 4.4. French group's accuracy in rejecting <i>wh</i> -island and complex NP violations	44
Table 4.5. Mean choice of answers corresponding to the scope of a matrix <i>wh</i> -word...	46
Table 4.6. Parameterized uFs in English and Spanish.....	48
Table 4.7. Martin's mean scores according to category	49
Table 4.8. Gender vs. number agreement errors.....	49
Table 4.9. Agreement (Agr) in production by stem shape (% target-like).....	56
Table 4.10. ERP effects to morphological violations in L2 and native speakers	58
Table 4.11. Response accuracy and effects to <i>wh</i> -dependencies in L2 learners and native speakers	60
Table 5.1. L2 informants' background information and self-rated language proficiency	67
Table 5.2. L2 Informants' foreign language configurations and age starting the language(s).....	68
Table 5.3. Patty's plural marking in obligatory contexts	69
Table 5.4. Informants' NPT scores (maximal points: 50).....	70
Table 5.5. Individual informants' accuracy on <i>et</i> items and <i>en</i> items	72
Table 5.6. Parameterized uninterpretable features in the L2 and the L1s	75
Table 5.7. Types of DPs and illustrations in the DP production task.....	77
Table 5.8. Absolute accuracy, accuracy on D and <i>a</i> for each L2 participant	81
Table 5.9. Each L2 group's errors on adjectives	83
Table 5.10. Each L2 group's suppliance of /t/, /e/, \emptyset in obligatory contexts.....	83
Table 5.11. Each L2 group's suppliance of <i>den</i> , <i>det</i> , <i>de</i> in obligatory contexts.....	84
Table 5.12. Each L2 group's errors on definite D	84

Table 5.13. Distribution of errors in each L2 group (%).....	86
Table 5.14. Rate of gender inconsistency at the individual and group level.....	87
Table 5.15. Individual G1 participants' gender alternation pattern.....	89
Table 5.16. Individual G2 participants' gender alternation pattern.....	90
Table 5.17. Phonetic forms of inflections on D, ADJ, and N in the Norwegian double definiteness construction	94
Table 6.1. Sample stimuli in the auditory naming task (mismatched features in bold form)	107
Table 6.2. Mean word frequency of the target nouns in each cell.....	108
Table 6.3. Mean syllable length of the target nouns in each cell	108
Table 6.4. Mean word durations in the recording (in milliseconds).....	109
Table 6.5. Control group's Mean RTs in each cell (RT in ms)	111
Table 6.6. The sub-control group's mean RTs in each cell (RT in ms)	113
Table 6.7. The L2 informants' mean RTs in each cell (RT in ms).....	114
Table 6.8. G1 L2 participants' mean RTs in each cell (RT in ms).....	116
Table 6.9. G2 L2 participants' mean RTs in each cell (RT in ms).....	116

Abbreviation list

Acc.: accusative Case
ADJ: adjective
ATB: across-the-board
C: complementizer
C-I: conceptual-intentional (interface)
CL: classifier
COM: common gender
D: prenominal determiner
DEF: definiteness
DM: Distributed Morphology
DP: determiner phrase
EEG: electroencephalogram
ERP: event-related brain potential
Exp: expression
F: feminine gender
FTFA: full transfer full access hypothesis
FFFH: failed functional features hypothesis
FL: the faculty of language
GB: Government and Binding theory
GEN: gender
iF: interpretable feature
IH: interpretability hypothesis
INDEF: indefiniteness
SLA: second language acquisition
LI: lexical item
L1: first language
L2: second language
MASC: masculine gender
MP: the Minimalist Program
ms: millisecond
MSIH: missing surface inflection hypothesis
N: noun
NEUT: neuter gender
NOM: nominative Case
NS: native speaker
NNS: nonnative speaker
NUM: number
OEHS: onset of an empty-headed syllable
PERF: perfective aspect
PL: plural
P&P: Principles and Parameters theory
PPh: phonological phrase
PTH: prosodic transfer hypothesis
PWd: prosodic word
Q: question

RDH: representational deficit hypothesis
RRC: restrictive relative clause
RT: reaction time
SG: singular
SMT: strong minimalist thesis
SM: sensory-motor (interface)
SSH: shallow structure hypothesis
t: trace
T: tense
Top: topic
uF: uninterpretable feature
UG: universal grammar
w: weak (inflection)

1 Overview

1.1 Introduction

A fundamental concern of second language acquisition (SLA) research is whether or not adult second language learners can ultimately reach nativelike attainment. Available SLA literature has shown that only in rare cases do adult interlanguage grammars fully converge with native grammars; by contrast, it has often been attested that interlanguage grammars diverge from native grammars in some important aspects, despite comprehensive exposure to the target language. Accounting for divergent outcomes of native speakers (NS) and nonnative speakers (NNS) thus has been taken as the central goal for the generative approach to SLA, the research paradigm within which the current study is conducted (see Hawkins 2001a, for the significance of explaining NS-NNS differences). It also constitutes the main goal of the present study.

Early attempts to account for NS-NNS differences have been centered on testing whether there is a critical period in language acquisition. The so-called *Critical Period Hypothesis* (CPH) was first formulated by Lenneberg (1967). The central idea of the CPH is that there is a critical period (the time span of which was initially labeled between age two and puberty) in (first and second) language acquisition; after this critical period, human capacity for learning languages declines with maturation, and nativelike linguistic competence can no longer be achieved. For researchers working within the Principles and Parameters (P&P) framework, a challenge has been to find out whether a critical period affects all linguistic domains, or none, or only certain subparts of Universal Grammar (UG). So far evidence from existing SLA studies unanimously points to no age effects on *UG principles*. Numerous researchers have demonstrated, typically through ‘poverty-of-the-stimulus’ phenomena,¹ that post-critical period second language (L2) learners still have access to UG principles (e.g. Kanno 1997, 1998; Pérez-Leroux and Glass 1999; Lozano 2003; Hawkins, *et al.* 2006, among others). For instance, in a recent study, Hawkins *et al.* (2006) reported that L2 learners showed

¹ The poverty-of-the-stimulus argument first appeared in Chomsky (1980:34), which claims that the primary linguistic data children are exposed to are impoverished with respect to their ultimate syntactic knowledge. Based on this argument, Chomsky posits that some aspects of linguistic knowledge are innate. The poverty-of-the-stimulus phenomena have often been taken as “the hallmark for the testimony for the notion that language is inherently innate in humans” (Ritter 2001:1). See Thomas (2002) for a historical account about the emergence of the poverty-of-the-stimulus argument.

sensitivity to the distribution of ‘fake’ reflexives in English, correctly accepting (a) examples and rejecting (b) examples below:

- (1.1) a. Calling for her dog, Sue shouted *herself* hoarse.
 * b. Calling for her dog, Sue shout hoarse.
- (1.2) a. Before Brian could smooth it off, the plaster dried hard.
 *b. Before Brian could smooth it off, the plaster dried *itself* hard.

Hawkins *et al.* (2006) argue that the distribution of fake reflexives is governed by syntactic constraints on the realization of argument structure, which is related to a Case requirement. Specifically, adjectives used as resultative complements (*hoarse* and *hard*) require “an appropriately positioned specifier” (Hale and Keyser 2002:5). Unaccusative verbs like *dry* contrast with unergative verbs like *shout* in the availability of such a specifier position in their respective argument structure. Such a specifier position is available in the argument structure underlying unaccusative verbs, hence no further argument need be projected in the derivation of the sentence; it is absent in the argument structure underlying unergative verbs, hence the resultative complement must project one, realized in English as a ‘self’ form. These constraints are invariant across languages that allow resultative constructions, hence are assumed to be a universal principle. The distribution of fake reflexives in English meets the three conditions designated by White (2003:23) for ‘poverty-of-the-stimulus’ instance in L2: (a) the property in question should not be available in the L1; (b) it must be underdetermined by input; (c) it should not be known through instruction or other metalinguistic means. Thus the L2 learners’ knowledge of the distribution of fake reflexives in English demonstrates that adult L2 learners indeed have access to UG principles. This entails that UG principles are not subject to a critical period.

As to the correlation between age effects and parameters, no consensus has been reached yet. L2 researchers are still debating whether or not UG-based parameters are affected by a critical period. Two opposing positions emerge in the current SLA literature. One position, represented by proponents of *Failed Functional Features Hypothesis (FFFH)*, e.g. Tsimpli & Roussou 1991; Hawkins and Chan 1997; Hawkins 2001b; 2004; Franceschina 2001; Hawkins and Liszka 2003; Tsimpli 2003), claims that (at least part of) parameterized properties are subject to a critical period, leading to representation deficits in L2 interlanguage grammars (hence it is also referred to as Representational Deficit Hypothesis (RDH)). The other position, represented by L2 researchers espousing *Full Transfer Full Access Hypothesis (FTFA)*, e.g. Schwartz and Sprouse 1996; White 2003), maintains that new parameters can be reset to settings of the target language even after a critical period, and that L2 speakers have the same underlying grammatical representations as native speakers. What is vulnerable to maturational constraints is the more peripheral components of the language faculty, such as the morphological module (termed the Missing Surface Inflection Hypothesis (MSIH), see for example, Lardiere 1998a, 1998b, 2000, 2006a; Prévost and White 2000) and the phonological module (termed the Prosodic Transfer Hypothesis (PTH), see for example, Goad *et al.* 2003; Goad and White 2004; Goad and White 2006). These different approaches to explaining NS-NNS divergence will be evaluated in the current study in relation to the L2 data from the experiments.

My study also explores another fundamental question concerning L2 ultimate attainment. This question regards whether endstate L2 speakers can attain nativelike processing capacity. This issue has not received much attention in the previous research on L2 ultimate attainment, probably because processing performance and grammatical competence have been viewed as separate subjects and belonging to different realms of research. This is a misconception. As pointed out by Hyltenstam and Abrahamsson (2003:576), it is a ‘problem’ that L2 researchers doing ultimate attainment studies tend to equate ‘language’ with ‘*grammatical competence*’. They further remark that “the domain for maturational constraints in general is the human capacity for language both at the level of knowledge and at the level of processing” (2003: 577). I fully agree with them on this point, and believe that our explanation for L1-L2 divergence will never be adequate if we do not link second language acquisition studies and psycholinguistic experimentation. In the present study, I follow the view of Hyltenstam and Abrahamsson and adopt a comprehensively defined notion of linguistic competence, including both grammatical competence and processing capacity. Investigating maturational constraints both at the level of knowledge and at the level of processing constitutes the primary goal of the current study.

The target language in this study is Norwegian, a language, together with Swedish, Danish, Icelandic, and Faroese, known as a Scandinavian language. To locate Norwegian in an upper branch of the language family tree, we find that it belongs to the Germanic language family, which makes it a close kin to English. But in fact Norwegian and English differ in many important aspects, including the syntactic patterning and the main inflectional categories, especially in the nominal domain. As we will discuss at length in Chapter 2, the Norwegian language is characterized by a rather complex inflectional system in its nominal domain, and there is agreement between the various constituents within nominal determiner phrases (DP). Norwegian DP internal agreement will be the focus in the current endeavor to investigate L2 grammatical knowledge and processing capacity. The topic was chosen because it has been shown that L2 speakers often have persistent difficulty with agreement morphologies (e.g. Long 1997; Lardiere 1998a, 1998b, 2006b; White 2002; Hawkins & Franceschina 2003). The potential area where adult L2 acquisition failure is likely to occur will make a good testing point for the various theoretical constructs seeking to explain L2-L1 divergence.

Acknowledging that “the most fruitful way to research maturational constraints is to focus explicitly on ultimate L2 learning” (Hyltenstam and Abrahamsson 2003), the current study investigates L2 knowledge and L2 processing capacity with respect to Norwegian DP internal agreement by examining the eventual outcomes of three groups of L2 learners. Since these informants used L2 Norwegian, along with their L1s, on a daily basis for a sufficient period of time (see Section 2 of Chapter 3 for details), they were virtually L2 speakers (or more generally L2 users), rather than L2 learners. The L2 speakers were selected from typologically different L1 backgrounds: Italian/Spanish, English, and Chinese. Syntactically, as we will discuss in some detail later (in Chapter 3), the three source languages differ parametrically from the target language with respect to one, two, or three of the uninterpretable features relevant for DP internal agreement. Thus, investigating the ultimate attainment of the three groups of L2 speakers will allow us to explore “the role of the L1 in delimiting what is ultimately attainable in the L2” (Lardiere 2006c:240).

It ought to be pointed out that the Norwegian language has rarely been the subject for generative second language acquisition research, contrasting its significant contributions to generative syntactic theories.² This seems to be a general trend for all Scandinavian languages, as can be seen from the following remark by Josefsson and Platzack (2004:1):

Although the Scandinavian languages have played an important role as research object in the development of generative grammar during the last twenty years, they have not been a prominent subject for studies of language acquisition within this research paradigm (Josefsson and Platzack 2004:1).

This situation is especially true for Norwegian. To my knowledge, the current study will virtually be the first doctoral dissertation on the acquisition of Norwegian as a second language embedded in the Principles and Parameters framework. It thus represents the initial efforts towards narrowing the gap of the contributions Norwegian has made to syntactic theories and to empirical SLA research. Through this study, I wish to demonstrate that the Norwegian language offers some special opportunities for the investigation of second language acquisition, especially regarding L2 ultimate attainment. Indeed L2 research requires the pairing of diverse source and target languages in order to gain a more comprehensive understanding about the nature of SLA, and about the nature of “the complex system of knowledge that has actually been attained” (Chomsky 1986).

1.2 Research questions

The present study addresses three main research questions with a view to investigating L2 grammatical competence and processing capacity. L2 knowledge with respect to Norwegian DP internal agreement is firstly investigated through an on-line DP production task, as on-line tasks are generally believed to be reliable measures of implicit linguistic knowledge (see for example, Ellis 2005). In addition, L2 learners’ knowledge about the Norwegian gender system is investigated separately, using a gender assignment task, as a gender feature is inherent in Norwegian nouns (contrasting, for example, number, which is a derived feature). These two experiments (which will also be referred to as L2 production tasks) together address the research question RQ1:

- RQ1: Can L2 speakers attain nativelike success in grammatical competence with respect to Norwegian DP internal agreement? If not, which of the current proposals (the FFFH/RDH vs. the MSIH or the PTH) is most compatible with the L2 data?

L2 processing will be tested using an auditory naming technique. As no previous study has been conducted on native speakers’ processing of Norwegian DP internal agreement, L1 processing performance will be investigated along with L2 processing. The auditory naming task can presumably reveal to us whether L1/L2 speakers can perceive the (concordant and discordant) gender/number/definiteness agreement cues

² There is ample research related to the syntactic structure of nominal determiner phrases alone, from the perspective of Norwegian or Scandinavian languages in general, see among others Taraldsen 1990; Svenonius 1993; Delsing 1993; Julien 2003, 2005.

available on the determiners (hence it will also be referred as the L2 perception task). This experiment examines research question RQ2:

- RQ2: Do endstate L2 speakers process Norwegian DP internal agreement the same way as the natives do?

Results from the two production experiments can give valuable insights on designing research questions for the subsequent study on processing performance. If some L2 speakers are found to perform in a nativelike way in the on-line production task, it is of interest to compare the processing behaviors of nativelike performers (those who presumably have attained nativelike grammatical competence) and non-nativelike performers (those who presumably have not attained nativelike grammatical competence) and see in what way L2 processing behaviors are related to L2 grammatical competence. Exploring the interrelation between L2 production and L2 perception gives rise to the third research question RQ3:

- RQ3: Is L2 speakers' nativelike (or non-nativelike) processing directly related to their L2 grammatical competence? In other words, does perception parallel production?

I am not aware of any previous research on ultimate attainment (e.g. Long 1997; Lardiere 1998a, 1998b, 2006b; Franceschina 2001; White 2002) addressing RQ3. This current study is probably the first to investigate the interrelation between L2 speakers' grammatical competence and syntactic processing capacity. At present, all the three research questions are of a general nature; they will be developed and refined in the subsequent chapters where the experiments are reported.

1.3 Thesis outline

This chapter briefly presents the research contexts and sets the background for the three main research questions in the current study. The remaining chapters are structured as follows. Chapter 2 gives a descriptive overview of the morphosyntax in the nominal domains of the target language and the source languages. As this study deals with the acquisition and processing of DP internal agreement, this chapter gives a comparative description of how DP internal agreement, if any, is realized in the languages under discussion. In Chapter 3, I present the analyses of DP structures of the target language and the source languages, based on the DP model proposed by Julien (2003, 2005). Parametric differences between the target and the source languages with respect to uninterpretable features, which are presumably susceptible to maturational constraints, are discussed at some length. As the current study and Julien's model are both conducted within the Minimalist framework, this chapter also includes in its initial section a brief introduction to the Minimalist Program, in particular its basic concepts and major claims, which are important for understanding the theoretical issues under investigation in this study. Chapter 4 is a review of SLA theories, including both theories on grammatical competence and theories on L2 processing. For the former part (i.e. theories on grammatical competence), the focus is on what I take to be the two competing SLA models within the generative framework, namely, the Full Transfer Full

Access hypothesis (FTFA) and the Failed Functional Features Hypothesis (FFFH). I will show that the two models represent the two opposing positions regarding age effects on parameter resetting, and as such are especially relevant to studies of L2 ultimate attainment. Their different views on maturational constraints and different approaches accounting for NS-NNS divergence will be presented and assessed. For the latter part (i.e. theories on L2 processing), a L2 processing theory known as the Shallow Structure Hypothesis (SSH, Clahsen and Felser 2006a, 2006b) will be outlined. In Chapter 5, I start with a detailed description of the L2 participants, as well as their scores achieved on the Norwegian proficiency test, then go on to report on two experiments which aim at ‘tapping’ L2 grammatical competence. The first experiment is an off-line gender-assignment task, where participants were asked to assign the correct gender to the Norwegian nouns; the second experiment is an on-line DP production task, through which Norwegian DPs containing adjectives were elicited. The results will be subsequently presented, and the extent to which the FTFA and the FFFH models can satisfactorily account for the empirical data will be discussed. The non-target-like data will be used to evaluate the various approaches (namely, the FFFH/RDH, the MSIH, and the PTH) to explaining NS-NNS divergence. Chapter 6 reports on an experiment exploring L1 and L2 processing behaviors. First an overview of the relevant psycholinguistic studies is provided, followed by a discussion of agreement processing mechanisms and a proposal for an agreement processing model. L1 and L2 results are subsequently presented and compared. A comparison is also made between the three L2 groups, and between the nativelike performers and non-nativelike performers in the on-line production task. This chapter ends with a discussion of L1 and L2 results in relation to current processing theories and implications of the finding for the SSH. Finally Chapter 7 provides answers to the research questions, summarizes the contributions of the thesis, and points out directions for future work.

2 *Morphosyntax of DPs in the target and the source languages*

2.1 Introduction

In this Chapter, I will deal with the structural aspect of the determiner phrases of the target language, Norwegian, and of the informants' first languages, namely, English, Italian/Spanish, and Chinese.³ Languages often differ as to what grammatical categories they encode, and as to the way a certain category is encoded. Among the three source languages, only Italian/Spanish has inflectional categories of gender, number, and definiteness in its nominal domains, as in the target language. Neither English nor Chinese has a gender system. While English encodes both number and definiteness, Chinese, as an isolating language, generally lacks gender, number, and definiteness distinctions, though suffix *-men* is arguably considered as a plural marker in some literature. The aim of this Chapter is to give a descriptive overview of the morphosyntax in the nominal domain of the languages under discussion. After a brief comparison of the nominal systems of the four languages in Section 2.2, I look at the similarities and differences between the Norwegian and Italian/Spanish gender and number systems and present a detailed description of how DP internal agreement is realized in those two languages in Section 2.3. A summary follows in Section 2.4.

2.2 A comparison of DPs of the target and source languages

In this section, I will make a general comparison of DPs in the target and source languages. Specifically, I will discuss the following three points: 1) whether a certain grammatical category has overt morphological reflexes in the DPs of L2 Norwegian and the source languages; 2) how a specific grammatical category is encoded, and 3) whether the languages exhibit DP internal agreement in terms of gender, number, or definiteness. This is followed by a discussion of whether there is a number agreement in Chinese and whether Chinese has number morphology at all.

³ In the present study, the term Chinese refers exclusively to *Mandarin Chinese*. Only the properties of Mandarin Chinese are discussed in this dissertation, and the Chinese informants in this study all acquired Mandarin Chinese as their native language.

As far as the nominal domain is concerned, the target language, Norwegian, has a relatively complex inflectional system, which encodes gender, number, and definiteness. And gender, number, and definiteness agreement (concord) is mandatory between nouns and determiners, nouns and adjective modifiers.⁴ This is illustrated in (2.1) below.⁵

- (2.1) a. et fin-t land
 D-NEUT.SG.INDEF nice- NEUT SG.INDEF country-NEUT.SG.INDEF
 ‘a fine country’
- b. de fin-e land-ene
 D-PL.DEF nice- DEF country-NEUT.PL.DEF
 ‘the nice countries’
- c. fin-e bil-er
 nice-PL car-COM. PL.INDEF
 ‘nice cars’

Note specially that in (2.1b) there is a co-occurrence of a free, prenominal determiner (*de*) and a suffixed determiner (*-ene*) on the noun. This construction is traditionally called a “double definiteness” construction. “Double definiteness” is a unique feature of the Norwegian language, and will be elaborated in the subsequent section.

Among the source languages, Italian/Spanish is the only language that has gender, number and definiteness inflections, and gender and number agreement must be marked on determiners, adjectives, and all coreferential pronouns. But unlike Norwegian, there is no definiteness agreement in Italian/Spanish. Some Italian examples are provided below to illustrate this point.

- (2.2) a. un libro carino
 D-MASC.SG. INDEF book-MASC.SG nice-MASC.SG.
 ‘a nice book’
- b. il libro carino
 D-MASC.SG.DEF book- MASC.SG nice- MASC.SG
 ‘the nice book’
- c. libri carini
 book- MASC.PL nice- MASC.PL
 ‘nice books’

English has a relatively simple nominal system, with inflections for number (-s as plural marker on the nouns; singular number is unmarked, and therefore not visible on

⁴ In addition, gender and number agreement is found between subjects and predicative adjectives. But this agreement relation is not of concern in the present study.

⁵ The following symbols are used for the annotation throughout the paper. *D* stands for prenominal determiners; *NEUT* for neuter gender, *COM* for common gender, *MASC* for masculine gender, *F* for feminine gender; *SG* for singular number, *PL* for plural number; *DEF* for definite, *INDEF* for indefinite.

the nouns). Moreover, English has number agreement between nouns and demonstratives, but not between definite articles and nouns, nor between adjectives and nouns, as shown in the following set of examples below:

- (2.3) a. the nice cup the nice cups
 b. this/that nice cup these/those nice cups

In comparison to most European languages, gender plays a rather small part in the grammar of English. With a few exceptions (e.g. *actor* vs. *actress*; *waiter* vs. *waitress*), English nouns generally do not encode a gender distinction. Accordingly, English does not exhibit gender agreement. Nonetheless the third person singular pronouns require the distinction between masculine *he*, feminine *she*, and neuter *it*, and there is gender agreement between nouns and their pronouns.

Among the languages in question, Mandarin Chinese is the only non-Indo-European language. It is known as an isolating language, a language without much inflectional morphology. In the nominal domain, Mandarin Chinese is first and foremost characterized by a total absence of grammatical gender: there is even no gender distinction for third person singular pronouns in speech, where *ta* serves as translation equivalent of all three English pronouns, though in the orthographic system they are distinguished according to masculine 他, feminine 她, and neuter gender 它.

The second striking feature of Chinese DPs is lack of number and definiteness markers; there is no morphological clue showing whether a noun phrase is definite or indefinite, singular or plural. Instead, a ‘bare’ noun can appear in any argument position, the number and definiteness of its referent being unspecified, as in (2.4):

- (2.4) Ta qu na pingguo.
 he/she go get apple
 ‘He/She went to get an apple/the apple/apples/the apples.’

A bare noun such as *pingguo* ‘apple’ in the above sentence can be interpreted as either singular or plural, definite or indefinite, the context providing the necessary information. Indeed, Chierchia (1998) proposes that Chinese NPs are all mass nouns in nature; mass nouns are ‘argumental’, and can be merged directly in argument positions in syntactic expressions without the need for a determiner. This is in contrast with Spanish/Italian (presumably Norwegian and English as well), whose NPs are ‘predicative’, hence can never be merged directly in argument positions, but require a determiner to license them as arguments.

Chinese is also known as having a rich classifier system. Classifiers “classify” referents of the nouns into a certain category based on some salient features of the objects that the nouns denote. For example, the classifier *ben* is used for books and magazines, whereas *zhang* is used for rectangular, flat objects such as tables, beds, maps. If a noun does not have its own specific classifier, then it takes *ge*, which is used as a general classifier. Syntactically, the classifiers have been argued to have “an individualizing or singularizing function” (Chen & Sybesma 1999), through which nouns become “syntactically visible and countable” (Doetjies, 1996). Hence a classifier is obligatory when the number of the individuals that a noun denotes is indicated, as we can see from the grammaticality contrast below:

- (2.5) a. san ge xuesheng b. * san xuesheng
 three CL student three student
 ‘three students’ ‘three students’

With respect to the number feature in Mandarin Chinese, there is a great controversy. Traditionally, the suffixes *-men* and *-xie* have been considered plural markers. However, nowadays there has been no general consensus on the status of the two suffixes, due to their distinct behavior from the plural marker in European languages and their co-occurrence constraints.

Unlike the English plural morphology *-s*, *-men* cannot be attached to all countable nouns; its distribution is restricted to animate nouns and personal pronouns. In addition, *-men* cannot occur with numerals, whether or not a classifier is present. This can be observed from the ungrammaticality of the following expressions:

- (2.6) a. * san ge xuesheng-*men* b. * san xuesheng-*men*
 three CL student three student
 Intended meaning: ‘three students’

As for *-xie*, a number of linguists (see for example, Lü 1980) consider it the plural marker for demonstratives *zhe* ‘this’ and *na* ‘that’, forming *zhe-xie* ‘these’ and *na-xie* ‘those’; others (e.g. Li & Thompson 1989; Chen & Sybesma 1999) treat it as an indefinite plural classifier. The issue which is of concern here is whether there is number agreement between the demonstratives and nouns in Chinese, as seemingly an agreement relation exists in the following Chinese noun phrases (for expository purposes, I temporarily use PL for both *-men* and *-xie* in the notation):

- (2.7) a. zhe -*xie* xuesheng(-*men*) b. na-*xie* xuesheng (-*men*)
 this-PL student-PL that-PL student-PL
 ‘these students’ ‘those students’

If it is indeed the case that *-xie* marks the plurality of the demonstratives, and *-men* the nouns, Chinese should have the same number agreement relation as English, the only difference being that the number suffix *-men* in Chinese is optional, while *-s* in English is obligatory. Indeed this is the claim made by Kuroda (1988): number agreement is obligatory in English, optional in Chinese.

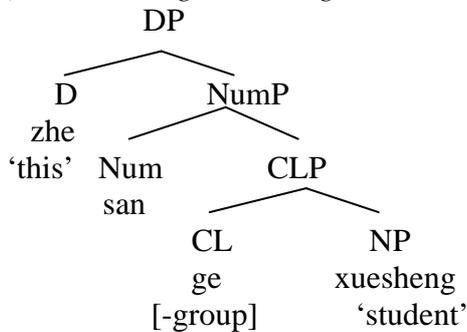
But any claim for *-men* and *-xie* as plurality markers for nouns and demonstratives, respectively, would have to explain why neither *-men* nor *-xie* can co-occur with a numeral (>1). The co-occurrence constraint of *-xie* is illustrated by the contrast below (Also see examples in (2.6) for the co-occurrence constraint of *-men*):

- (2.8) a. zhe san ge xuesheng b. * zhe-xie san ge xuesheng
 this three CL student this-PL three CL student-PL
 ‘these three students’ ‘these three students’

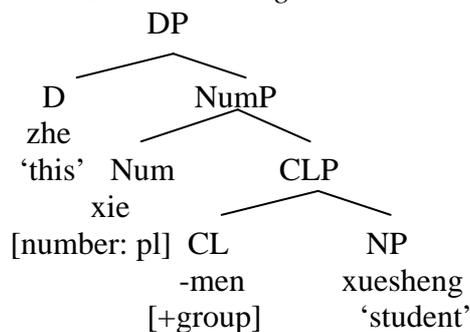
The examples in (2.8) suggest that the singular demonstrative can co-occur with numerals (>1), whereas the plural demonstrative cannot—a fact that is in reverse to those languages that display number agreement, such as English, Norwegian. To

account for this puzzle, I follow Yang (2004) in arguing that *-men* is a suffixal classifier instead of a plural marker, and that *xie* is a plural element under Num⁰. Under this account, the demonstrative phrases *zhe san ge xuesheng* ‘these three students’ and *zhexie xuesheng (-men)* ‘these students’ have a syntactic representation in (2.9a) and (2.9b), respectively:⁶

(2.9) a. *zhe san ge xuesheng* ‘these three students’



b. *zhexie xuesheng* ‘these students’



Yang assumes that the suffixal classifier *-men* has a [+group] feature, which gives rise to a group interpretation of an NP. She further argues that the lexical item *xie* is the overt realization of the plural feature under the Num head. Numerals themselves, however, do not have a plural feature. Since both *xie* and numerals are elements on the Num head, they compete for the same syntactic position, hence the two cannot co-occur. This accounts for the co-occurrence constraint between *xie* and numerals (>1). After NP merges with CL with a [+group] feature as illustrated in (2.9b), N head moves to CL to pick up the [+group] feature of *-men*. The next step when Num is merged, it must contain a [+pl] feature, otherwise there will be a feature clash between the [+group] feature on CL+N and the [-pl] feature on Num. If Num has a plural feature, under Yang’s assumption, the plural feature has to be realized as *xie*. On the other hand, if NP merges with CL with a [-group] feature, as illustrated in (2.9a), numerals which are without a plural feature are compatible with the [-group] feature on CL. Crucially, D in Chinese does not have a number feature, so in the above two cases, *zhe* ‘this’ is spelt out irrespective of the number features on the nouns.

In short, Yang proposes that 1) *-men* is a suffixed classifier, instead of a plural marker, 2) *xie* is a plural element under Num, instead of a plural suffix of

⁶ The syntactic structure of Chinese DP presented here is based on Yang (2004), though it is subject to modification in the subsequent chapter.

demonstratives, and 3) D in Chinese does not involve a number feature. Her proposal, in my view, nicely captures the co-occurrence constraints of *xie* and *-men* with numerals (>1). More relevant to my discussion is the conclusion that Chinese does not show D-N agreement in terms of number.

To summarize so far, we have discussed whether the languages in question have overt morphology in terms of gender, number, and definiteness, and whether they exhibited DP-internal agreement with respect to these features. As can be seen from Table 2.1, the target language, Norwegian, has overt morphological reflexes of number, gender and definiteness. It also has overt agreement between determiners, adjectives, and nouns in terms of number, gender and definiteness. Italian/Spanish share all the nominal properties with Norwegian except for definiteness agreement, while Chinese is distinct from Norwegian in that it has no overt morphological expression for all these nominal properties. English, like Chinese, has no gender or gender agreement in its nominal domain, but it encodes number and definiteness, and it exhibits overt number agreement between demonstratives and nouns.

Table 2.1. Inflectional categories and agreement in DPs of the L2 and the L1s

	Norwegian	Italian/Spanish	English	Chinese
Number	+	+	+	-
Number agreement	+	+	+	-
Gender	+	+	-	-
Gender agreement	+	+	-	-
Definiteness (DEF)	+	+	+	-
DEF agreement	+	-	-	-

Key: + Present in language; - absent in language

2.3 Gender and number systems of Norwegian and Italian/Spanish

As has been mentioned before, Italian/Spanish is the only source language, like the target language Norwegian, which has gender and number agreement. But as members of different language families (namely, Germanic and Romance), Norwegian and Italian/Spanish have structurally distinct gender and number systems. Since the structure of gender and number systems in the first language presumably has some effect on L2 acquisition and perception, this section will be devoted to a detailed discussion of the similarities and differences between the gender and the number systems in the two languages. The gender and the number systems in English and Chinese are much simpler, and have been touched upon in the previous section, so they are not subjected to further discussion in this section.

2.3.1 Gender and number in Norwegian

Norwegian is a language with a more formal gender system than a semantic gender system (see Corbett 1991, for a distinction between formal and semantic gender systems). Generally, the gender of the nouns bears no relation with the sex of nouns' referents. For example, *lærer* 'teacher' is assigned a masculine gender, irrespective of

the fact that it refers to a male or a female teacher.⁷ In many Norwegian dialects, there are three genders: neuter, masculine and feminine. The three genders are not evenly distributed, though. According to the *Oslo Corpus of Tagged Norwegian Text* (website address: <http://www.hf.uio.no/tekstlab/frekvensordlister/index.html>), the frequency of occurrences for the three indefinite articles in magazines and newspapers are 126870 for *en* (masculine indefinite), 57635 for *et* (neuter indefinite), and 957 for *ei* (feminine indefinite). This means masculine nouns make up 68% of all the indefinite singular nouns in the corpus, neuter nouns 31%, and feminine nouns only 1%. The small proportion of feminine nouns can be (partly, at least) attributed to the fact that in Bokmål Norwegian, masculine and feminine genders can be collapsed into a common gender (*en*), which is in the same form as the masculine gender. For example, *jente* ‘girl’ is a feminine noun (e.g. *ei jente*) in the three way gender system. In the two way system, it has a common gender (i.e. *en jente*) instead. Actually in the three-way gender system, the distinction between masculine and feminine gender can only be visible on singular indefinite articles, and optionally on the suffixed definite article on nouns; the agreement inflections on the prenominal definite articles and adjectives are exactly the same for both genders, as shown in the two sets of examples in (2.10), illustrated by a feminine noun *jente* ‘girl’ and a masculine noun *bil* ‘car’. In order to simplify assumptions, this paper adopts the binary gender system of Norwegian. That is to say, Norwegian nouns distinguish a neuter gender from a common (non-neuter) gender. Common gender, which is over twice as frequent as neuter gender, is considered as the default or the unmarked form, whereas neuter gender is the marked form.

- (2.10) a. *ei fin jente* *den fine jenta/en* *de fine jentene*
 ‘a fine girl’ ‘the fine girl’ ‘the fine girls’
- b. *en fin bil* *den fine bilen* *de fine bilene*
 ‘a fine car’ ‘the fine car’ ‘the fine cars’

Gender assignment often seems to be largely arbitrary in Norwegian. In addition, phonological forms of nouns give no cue to their gender.⁸ For example, *ball* ‘ball’ is assigned a common gender, while *tall* ‘number’ a neuter one; *brann* ‘fire’ is assigned a common gender, while *våpen* ‘weapon’ a neuter one; *kirke* ‘church’ is assigned a common gender, while *ønske* ‘wish’ a neuter one.

Cues to gender classification of nouns are exhibited in determiners, attributive adjectives, and various other elements that co-occur with the head noun. In singular indefinite neuter context, gender (and number) agreement has an overt phonological realization both on the determiner and on the attributive adjective, while in the corresponding common gender, it has a phonological realization only on the determiner, but not on the attributive adjective (the null morpheme is marked with a zero symbol ‘ \emptyset ’). This is illustrated in (2.11) below.

⁷ There is an antique form *lærerinne* for female teachers, but this form is rarely used in modern Norwegian.

⁸ Certain derivational suffixes, however, may provide some morphological cues to gender. Suffixes, such as *-sjon*, *-else*, *-ning*, *-het*, *-dom*, *-nad*, are associated with common gender, whereas *-ment*, *-eri*, *-em*, *-tek* are associated with neuter gender.

- (2.11) a. et fin-t land
 D-NEUT.SG.INDEF nice- NEUT country-NEUT.SG.INDEF
 ‘a fine country’
- b. en fin bil
 D-COM.SG.INDEF nice- COM(Ø) car-COM.SG.INDEF
 ‘a nice car’

Also note that the realization of *-t*, which marks neuter indefinite on adjectives, is rather restricted. Generally adjectives ending in 1) *-t* or *-d*, and 2) in a derivational suffix, remain invariant. Actually such adjectives are very pervasive, and are actively involved in everyday life. Adjectives from the first category include *lett* ‘easy’, *fast* ‘steady’, *glad* ‘happy’, *opptatt* ‘busy’; from the second category, *vanskelig* ‘difficult’, *tidlig* ‘early’, *viktig* ‘important’, *alvorlig* ‘serious’, *norsk* ‘Norwegian’, *amerikansk* ‘American’, to name just a few. This means in many cases this gender agreement is not visible on the adjectives.

More complication is involved in Norwegian definite determiner phrases. Unlike indefinite DPs, where the indefinite determiner is a prenominal free morpheme, the definite determiner is realized as a suffixed bound morpheme, which encodes gender and number information as well. What is more intriguing, when an attributive adjective (or a numeral) is inserted in the definite DPs, it requires the co-occurrence of both a suffixed determiner and a free prenominal determiner. Hence, when modified, the Norwegian definite DPs exhibit a definiteness agreement between nouns and determiners, apart from gender and number agreement. Due to what appears to be double marking of definiteness, this construction is commonly referred to as a ‘double definiteness’ construction.⁹ In contrast to the ‘strong’ inflection in indefinite DPs, where adjectives show number and gender distinctions, the adjectives in the double definiteness construction take an invariant suffix *-e*. This invariant form is often referred to as weak inflection, and I use *w* to stand for it in annotation. The following are examples illustrating the agreement paradigm of Norwegian definite DPs:

- (2.12) a. land-et
 country- NEUT.SG.DEF
 ‘the country’
- b. bil-en
 car-COM.SG. DEF
 ‘the car’
- c. det fin-e land-et
 D-NEUT.SG.DEF nice-*w* country-NEUT.SG.DEF
 ‘the nice country’
- d. den fin-e bil-en
 D-COM.SG.DEF nice-*w* car-COM.SG.DEF

⁹ Apart from Norwegian, the ‘double definiteness’ construction can be found in Swedish and Faroese, but not in other Scandinavian languages, such as Danish and Icelandic.

‘the nice car’

In addition, Norwegian demonstrative phrases also require double definiteness, irrespective of presence or absence of a modifier. Demonstratives are inflected for gender and number. Besides, they also distinguish on the proximity. Examples of demonstrative double definiteness are given below. Note that the distal forms of demonstratives resemble the free pronominal determiners.

(2.13) Distal

a. det (fin-e) land-et
that-NEUT.SG. (nice-w) country- NEUT.SG.DEF
‘that (nice) country’

b. den (fin-e) bil-en
that-COM.SG. (nice-w) car-COM.SG.DEF
‘that (nice) car’

(2.14) Proximal

a. dette (fin-e) land-et
this-NEUT.SG. (nice-w) country- NEUT.SG.DEF
‘this (nice) country’

b. denne (fin-e) bil-en
this-COM.SG. (nice-w) car-COM.SG.DEF
‘this (nice) car’

It should be pointed out that this ‘double marking’ of definiteness in modified Norwegian DPs is not obligatorily required in all cases. ‘Single definiteness’, where either the suffixed determiner or the pronominal determiner is left out, is either optional or obligatory, depending on the contexts and various other factors. The pronominal determiner is often left out in nominals which are deictic (cf. example 2.15a), vocative (cf. example 2.15b), depictive (cf. example 2.15c), or are associated with a proper name (cf. example 2.15d) (Julien 2005: 30-34; Delsing 1993:116-120).

(2.15) a. Ta (den) ny-e bil-en. (adapted from Delsing 1993:118)

take (the) new-w car-COM.SG.DEF
‘Take the new car.’

b. Vet du ikke det, (*den) stor-e jent-en! (Julien 2005:32)

know you not that (D- COM.SG.DEF) big-w girl-COM.SG. DEF
‘Don’t you know that, you big girl!’

c. Det var (*den) svart-e natt-en da hun kom. (ibid)

it was (D-COM.SG.DEF) black-w night-COM.SG.DEF when she came
‘It was dark night when she came.’

d. Vi besøkte Hvit-e Hus-et. (ibid)

we visited white-*w* house-NEUT.SG.DEF
 ‘We visited the White House.’

On the other hand, the suffixed determiner can be left out if the nominals get a non-specific reading (Julien 2005:35-39). Julien points out the specific/nonspecific contrast using the following examples:

(2.16) a. Jeg lik-er det hvit-e gull-et bedre.
 I like-PRES D-NEUT.SG.DEF white-*w* gold-NEUT.SG. DEF better
 ‘I like the white gold better.’

b. Olje kalle-s iblant ‘det sort-e gull’.
 oil call-PASS sometimes D-NEUT.SG DEF black-*w* gold
 ‘Oil is sometimes called ‘the black gold’.

The DP in example (2.16a), *det hvite gullet*, has a specific reading; whereas the DP in (2.16b), the one without the suffixed determiner on the noun, gets a non-specific reading. What is NOT mentioned in the literature is that the suffixed determiner can also be left out in a formal speech or written discourse, without any semantic effect (Åfarli 2005, personal communication). In a sense, the drop of the suffixed determiner in the ‘should-be’ double definiteness construction simply signals the formal genre of the discourse. This particular usage to a certain extent reflects the residue of Danish influence on the Norwegian language. Danish resembles Norwegian within DPs in all other aspects except that it does not have the ‘double definiteness’ construction, as one can observe from the following set of Danish DP examples:

- (2.17) a. et land
 D-NEUT.SG. country-NEUT.SG.
 ‘a country’
- b. land-et
 country- NEU.SG.DEF
 ‘the country’
- c. et fin-t land
 D-NEUT.SG. nice- NEUT country-NEUT.SG.
 ‘a fine country’
- d. det fin-e land
 D-NEUT.SG. nice-*w* country-NEUT.SG.
 ‘the nice country’

We see that in Danish, there is a pronominal determiner and no suffixed article in definite DPs containing an attributive adjective. The Bokmål Norwegian (which used to be called Dano-Norwegian) was actually developed on the basis of Danish, while integrating into it “specifically Norwegian elements from the colloquial speech of the

educated classes in urban areas (Askedal 1994).” It is thus natural that Norwegian retains some traces of Danish syntactic norms in its formal style of speech and writing.

With respect to number, Norwegian has singular (SG) and plural (PL) distinctions. Singular number is the unmarked form (or zero form), while plural is the marked form. Plural is generally expressed with a suffix, except for monosyllabic neuter nouns in indefinite context, which remain invariable in the plural (cf. example 2.18a). When the noun is plural, gender information is not visible on the plural suffixes, which are *-er* and *-ene* for indefinite and definite respectively. This is illustrated in the following set of examples:

- (2.18) a. (mange) land [neuter, monosyllabic]
 (many) country-PL(Ø).INDEF
 ‘many countries’
- b. land-ene
 country-PL. DEF
 ‘the countries’
- c. (mange) bil-er [common gender, monosyllabic]
 (many) car- PL.INDEF
 ‘many cars’
- d. bil-ene
 car- PL. DEF
 ‘the cars’

In addition, adjectives and pronominal determiners all have plural forms. The plural form of the pronominal determiner is *de*, irrespective of the gender of the following noun. Plural indefinite DPs do not require a pronominal determiner, or put another way, they require a null pronominal determiner (which is symbolized with a zero marker \emptyset). The plural form of adjectives takes an invariant suffix *-e*, both in definite and indefinite contexts. Modified plural Norwegian DPs are illustrated in (2.19):

- (2.19) a. \emptyset fin-e land
 nice-PL country-NEUT.PL (Ø).INDEF
 ‘nice countries’
- b. de fin-e land-ene
 D-PL.DEF nice-w country- NEUT.PL.DEF
 ‘the nice countries’
- c. fin-e bil-er
 nice-PL car-COM. PL.INDEF
 ‘nice cars’
- d. de fin-e bil-ene
 D- PL. DEF nice-w car-COM.PL.DEF
 ‘the nice cars’

To sum up, I have presented a rather comprehensive overview of Norwegian DPs and realizations of DP-internal agreement. Table 2.2 lists the regular inflectional paradigm of modified Norwegian DPs with all possible feature combinations.¹⁰

Table 2.2. Inflection paradigm of modified Norwegian DPs

FEATURE BUNDLES	D	ADJ	N
[SG][NEUT][INDEF]	et	-t	-∅
[SG][COM][INDEF]	en	-∅	-∅
[SG][NEUT][DEF]	det	-e	-et
[SG][COM][DEF]	den	-e	-en
[PL][INDEF]	∅	-e	-er
[PL][DEF]	de	-e	-ene

As we see, (the prenominal) D has altogether six (free) morphological forms; ADJ has three realizations: *-t* marks neuter singular indefinite; zero (*-∅*) marks common singular indefinite; *-e* marks the rest. N has five inflectional forms, zero (*-∅*), *-et*, *-en*, *-er*, and *-ene*. The irregularities of inflectional system and context-dependent omission of determiners in the double definiteness construction add further complexity to this picture. But the present study focuses on the regular inflection paradigm. The inflectional paradigm in Table 2.2 above thus provides an important reference point in assessing L2 learners' success with respect to DP internal agreement.

2.3.2 Gender and number in Italian/Spanish

As is the case for all Romance languages, Italian and Spanish have gender and number distinctions, and both languages have DP internal agreement in terms of gender and number. Italian and Spanish gender and number systems have a lot in common, so for the purpose of this study I consider them as one system. At the points where they differ from each other, I will explicitly point out.

Noun gender in Italian and Spanish falls into two categories: masculine and feminine. This is the first point where Norwegian and Italian/Spanish deviate. The two way gender distinction in Italian/Spanish, corresponding to masculine and feminine, do not map into the two way gender system of Norwegian, corresponding to common gender and neuter. For example, Italian nouns *mela* 'apple' and *palazzo* 'building' are feminine and masculine respectively, whereas their Norwegian counterparts, *eple* and *bygg*, are both of neuter gender. The second point where Italian/Spanish and Norwegian gender systems differ is that in Italian and Spanish, gender is to a large extent transparently marked on word endings. The general rule is that nouns ending in *-o* are masculine; nouns ending in *-a* are feminine. To use Bates *et al*'s (1996) metaphor, the noun 'wears its gender on its sleeve'. In contrast, the phonological forms of Norwegian nouns generally provide no cue to their gender. So learners of Norwegian cannot rely on word endings; rather they have to learn the gender on a word-by-word basis. Of course, there are nouns whose genders are not predictable on the basis of word endings in Italian/Spanish as well. For example, Italian nouns ending in *-e* can either be masculine

¹⁰ The inflection system presented here is based on the Bokmål variety of Norwegian. There are dialectal variations in this inflectional paradigm.

(e.g. *ristorante* ‘restaurant’) or feminine (e.g. *stazione* ‘station’). But such nouns only constitute a small number; the vast majority of nouns in Italian and Spanish ends in *-o* or *-a*, which is respectively associated with masculine and feminine. For purposes of the present study, it is important to point out that there is a morphology-gender correlation in Italian/Spanish nouns, whereas word endings of Norwegian nouns provide no reliable cue to their gender.

In Italian/Spanish, the gender inflection on adjectives generally follows the same pattern as the nouns. The adjective can be either pre-nominal or post-nominal. Its syntactic position, however, does not affect the agreement pattern, and is therefore not of concern in this paper. The determiners, definite or indefinite, are always prenominal, and the inclusion of attributive adjectives does not have any consequences on the determiners (in contrast to Norwegian, whose double definiteness construction results from an inclusion of an adjective/numeral in the definite context), as illustrated by the following Italian and Spanish examples:

(2.20) a. un libro carino
 D-MASC.SG.INDEF book-MASC.SG nice-MASC.SG.
 ‘a nice book’

b. una casa carina
 D-F.SG.INDEF house-F.SG nice-F.SG.
 ‘a nice house’

c. il libro carino
 D-MASC.SG.DEF book-MASC.SG nice-MASC.SG
 ‘the nice book’

d. la casa carina
 D-F.SG.DEF house-F.SG nice-F.SG
 ‘the nice house’ (Italian)

(2.21) a. un libro fino
 D-MASC.SG.INDEF book-MASC.SG nice-MASC.SG.
 ‘a nice book’

b. una casa fina
 D-F.SG.INDEF house-F.SG nice-F.SG
 ‘a nice house’

c. el libro fino
 D-MASC.SG.DEF book-MASC.SG nice-MASC.SG
 ‘the nice book’

d. la casa fina
 D-F.SG.DEF house-F.SG nice-F.SG
 ‘the nice house’ (Spanish)

From the above examples we can see that in the singular context, the agreement patterns in Italian and Spanish are very much alike. In the plural context, however, they deviate from each other, because the number systems in the two languages are rather different. In Spanish, plural forms of nouns, adjectives, determiners are marked consistently with the suffix *-s*, whereas the plural of Italian nouns is realized by inflections on the word final vowels: nouns ending in *-o* (Masc.) change the final *-o* to *-i*; nouns ending in *-a* (F.) change the final *-a* to *-e*. But one thing in common for both Italian and Spanish is that the gender distinctions are still visible in the plural (cf. Norwegian, where gender distinctions are not visible in the plural suffixes). The attributive adjectives and the definite determiners in Italian/Spanish generally take the same inflectional forms as their head nouns. The indefinite plural nouns in Italian, like in Norwegian, do not require the presence of a determiner. In Spanish, however, a determiner is obligatory in any contexts. Examples of Italian and Spanish agreement patterns in plural contexts are given below:

- (2.22) a. \emptyset libri carini
 book- MASC.PL nice- MASC.PL
 ‘nice books’
- b. \emptyset case carine
 house-F.PL nice- F.PL
 ‘nice houses’
- c. i libri carini
 D-MASC.PL.DEF book- MASC.PL nice- MASC.PL
 ‘the nice books’
- d. le case carine
 D-F.PL.DEF house- F.PL nice- F.PL
 ‘the nice houses’ (Italian)
- (2.23)
- a. unos libros finos
 D-MASC.PL.INDEF. book- MASC.PL nice- MASC.PL
 ‘nice books’
- b. unas casas finas
 D-F.PL.INDEF. house- F.PL nice- F.PL
 ‘nice houses’
- c. los libros finos
 D-MASC.PL.DEF book-MASC.PL nice- MASC.PL
 ‘the nice books’
- d. las casas finas
 D-F.PL.DEF house- F.PL nice- F.PL
 ‘the nice houses’ (Spanish)

2.4 Summary

In this Chapter I have presented a descriptive overview of the morphosyntax of DPs in the target and source languages. I will now recapitulate some important differences regarding determiner phrases of the languages under discussion. All the languages in question, except for Chinese, encode definiteness, but only Norwegian exhibits overt definiteness agreement. This definiteness agreement is most clearly manifested in the double definiteness construction, where there is a co-existence of a suffixed determiner and a free pronominal determiner. Both Norwegian and Italian/Spanish nouns have gender distinctions, but the two languages have distinct gender systems. Norwegian nouns distinguish neuter gender from common (non-neuter) gender. The common gender, which is over twice as frequent as the neuter gender, is assumed to be the default gender. Italian/Spanish also has a two-way (masculine vs. feminine) gender system, but Italian/Spanish gender does not usually map into Norwegian gender. Besides, most nouns are transparently marked for their gender by word endings in Italian/Spanish. In Norwegian, by contrast, gender is not predictable by morphological forms of the nouns. Gender plays a very small part in English, and almost no role in Mandarin Chinese. Italian/Spanish, like Norwegian, has DP-internal agreement in terms of Gender and number. English has number concord between demonstratives and nouns, but not between articles and nouns. Chinese, as I have argued, does not display any overt number agreement. These cross-linguistic observations will form the basis for assuming parametric variations in feature makeup of functional heads in the DP structures of the languages in question. This issue will be discussed and elaborated in the subsequent chapter.

Table 2.3 gives a set of examples illustrating DP internal agreements in Norwegian, Italian/Spanish and English.

Table 2.3. Illustrations of DP internal agreement in the L2 and L1s

		Norwegian		Italian/Spanish		English
		COM	NEUT	MASC	F	
SG	INDEF	en fin bil 'a nice car'	et fint land 'a nice country'	un libro carino <i>un libro fino</i> 'a nice book'	una casa carina <i>una casa fina</i> 'a nice house'	a nice book
	DEF	den fine bilen 'the nice car'	det fine landet 'the nice country'	il libro carino <i>el libro fino</i> 'the nice book'	la casa carina <i>la casa fina</i> 'the nice house'	the nice book
PL	INDEF	fine biler 'nice cars'	fine land 'nice countries'	libri carini <i>unos libros finos</i> 'nice books'	case carine <i>unas casas finas</i> 'nice houses'	nice books
	DEF	de fine bilene 'the nice cars'	de fine landene 'the nice countries'	i libri carini <i>los libros finos</i> 'the nice books'	le case carine <i>las casas finas</i> 'the nice houses'	the nice books

3 *Theoretical assumptions*

3.1 Introduction

Having given an overview of the crosslinguistic facts regarding the morphosyntax in the nominal determiner phrases of the target and the source languages in the previous chapter, I will present the analyses of DP structures of the languages under discussion in the present chapter. I am primarily interested in the parametric differences in feature makeup of functional heads in the DP structures of the languages in question. This study adopts as a theoretic framework the recent versions of the Minimalist Program (Chomsky 1999, 2000, 2001, forthcoming). Before presenting a minimalist account for DP structures of the target language Norwegian and the source languages, I will first introduce in Section 3.2 some basic concepts and sketch the major claims of the Minimalist Program (MP), which will become essential in the subsequent analysis of the syntactic structure of DPs. Following up on this, I present the syntactic structure of Norwegian DPs in Section 3.3, Italian/Spanish DPs in 3.4, and English DPs in 3.5. In Section 3.6, I make a brief note on the syntactic structure of Chinese DPs. All these syntactic analyses are based on the DP model proposed by Julien (2003, 2005). A summary of this chapter is spelt out in Section 3.7.

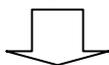
3.2 Some major claims of the MP

The MP, like the Government and Binding (GB) framework before it, adopts the Principles-and-Parameters (P&P) approach to natural languages and language acquisition. The general idea behind the P&P approach is that the course of language acquisition is determined by a biologically endowed innate faculty of language FL. This view has been popularly known as the *Innateness Hypothesis*. FL is “a component of the human brain dedicated to language” (Chomsky 2000). It is regarded as an ‘organ of the body,’ one of the many subcomponents of an organism that interact in its normal life. The theory of Universal Grammar (UG) is the theory of the initial state of FL. The P&P assumes that UG incorporates a set of universal grammatical principles and a series of parameters with binary settings. Principles are responsible for the underlying similarities, while parameters account for language diversity. Under the P&P approach, (first) language acquisition has been assumed to involve the task of setting a number of grammatical parameters to appropriate values in response to positive linguistic input.

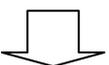
The P&P approach is first incarnated by the GB theory. Under the GB framework, languages could vary along many dimensions. For example, it has been assumed within the GB framework that Italian and English are parameterized with respect to the null subject parameter — Italian has a [+null subject] value, while English has a [-null subject] value (cf. Rizzi 1982; Radford 2004). It is also assumed that languages can be parameterized in terms of feature strength. Strong vs. weak C, for example, has been argued to be the reason underlying syntactic contrasts between V2 and non-V2 languages (cf. Travis 1984).

The MP was launched with a view to economizing the GB accounts (Hornstein, 2001). As far as parameters are concerned, feature strength and non-lexical parameters are not well motivated, and are eliminated in the MP. In light of the MP, a language is specified according to the following processes. First, UG provides a universal set of phonological, semantic-conceptual, and formal features F and syntactic operations that access F to generate expressions (Exp). But access to the universal set of features F is restricted, so in the next step, the language L makes a one-time selection of a subset $[F_L]$ of F . $[F_L]$ are the features that enter into L ; other features can be disregarded in the use of L . Finally, L makes a one-time operation that assembles elements of $[F_L]$ into a Lexicon, in which lexical items (LIs) are assembled. LIs are ‘signs’ expressing a pairing of phonological-semantic-formal features. Phonological features are relevant to phonological representation; semantic features are relevant to semantic representation; formal features fall into two categories: interpretable (iF) and uninterpretable (uF). This means that the features which get into derivations are ‘filtered’ from the universal set F in several successive steps, as illustrated in (3.1) below:

(3.1) **F**: Universal set of (phonological, semantic and formal) features



[F_L]: Subset of the F selected into the lexicon of a particular language



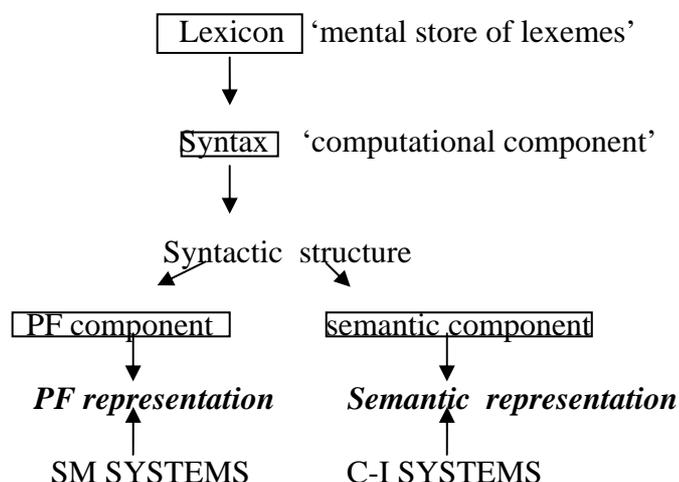
Lexicon: A set of lexical items constructed from $[F_L]$

Hence in the minimalist view, parametric variation is restricted to the lexicon, and the (first) language acquisition process consists of selection of (phonological, semantic and formal) features $[F_L]$ from universal feature set and assembling features from $[F_L]$ to construct lexical items (LIs) (cf. Chomsky 2000, p101.).

In essence, what motivated the MP is the requirement to minimize the theoretical apparatus used to describe language. In this spirit, Chomsky proposes the following methodological guidelines: Strong Minimalist Thesis (SMT), which assumes that language is an optimal solution to interface conditions, and Strong Uniformity Thesis (SUT), according to which languages are assumed to be uniform, with variety restricted to easily detectable properties of utterances. Given these methodological guidelines, only two interface levels are considered as “optimal properties”, hence retained within the MP: the sensory-motor interface (SM) and the conceptual-intentional interface (C-

I).¹¹ Accordingly, the internal organization of Grammar can be illustrated as in (3.2) below (based on Chris Wilder, lecture notes, spring 2006).

(3.2)



As has been discussed earlier, the lexicon consists of a set of lexical items (LIs). It is minimal in that only non-redundant information is encoded. Hence it is a set of lexemes, rather than phonological/morphosyntactic words. The lexical items are fed into the narrow syntax, where sentences are built up through a series of syntactic computations, thereby forming a syntactic structure. This syntactic structure then serves as input to the PF component and the semantic component, respectively. At this juncture, the PF component maps a syntactic structure into the corresponding PF representation; the semantic component maps it into the corresponding semantic representation. The semantic representation interfaces with C-I systems, and the PF representation with SM systems.¹²

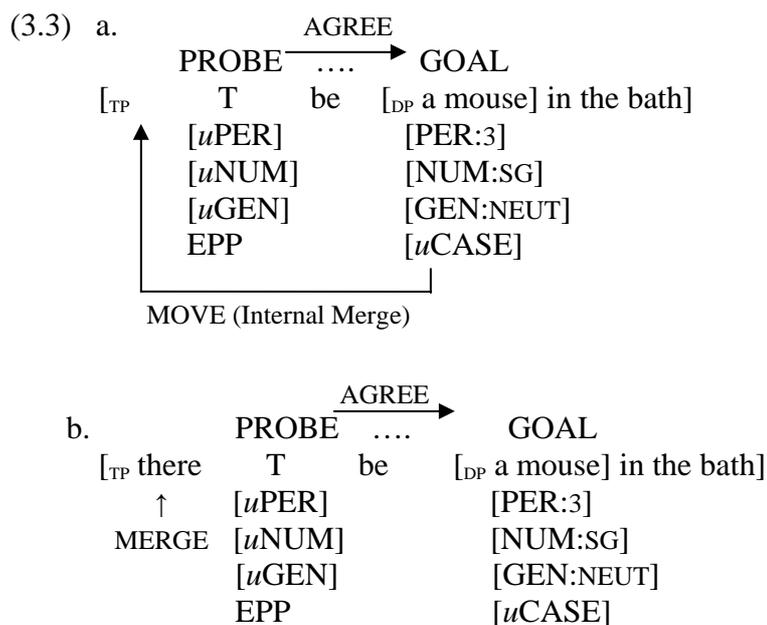
Given the Strong Minimalist Thesis (SMT), uFs, which enter into derivation unvalued, are not sanctioned at the C-I interface, and must be eliminated for the purpose of convergence. This is done via an operation called Agree, which establishes an agreement relation between the uFs and the matching iFs. Using Chomsky's terms, the head containing uFs is a Probe, seeking a matching Goal, that is, an element with iFs. The agreement relation between Probe and Goal must be satisfied locally, with the Goal in the C-command domain of the Probe. Once the agreement relation is established between the Probe and the Goal, feature valuation will take place via a Feature-copying operation, through which the feature values on the Goal are copied onto the Probe. If the Goal has an EPP feature, this agreement may require overt movement, in which case the Goal is forced to move to the Spec of the Probe head. However, agreement without movement is possible, in fact compulsory, if the EPP feature on the Goal can be

¹¹ Other levels postulated within the P&P framework such as d-structure, s-structure and Logical Form (LF) have been eliminated in the Minimalist spirit. However, in recent literature the term LF has often been used to refer to the C-I interface itself, as pointed out by Chomsky in footnote 11 of 'On Phases' (ms., downloadable from http://www.phon.ucl.ac.uk/home/hans/mrg/chomsky_onphases_1204.pdf). To avoid confusion, I chose to use the new terms (SM & C-I), while keeping the term LF in its original sense.

¹² Note that there are various proposals regarding levels of representation in the minimalist program. For an introduction to different models of representation, see Boeckx (2006:72-77).

satisfied by *Merge*. Merge combines two LIs. It is independently needed for any recursive system. Hence Merge “comes free”. Move is composed of Agree and (re)Merge, so it is more costly than Merge. Therefore, Merge is preferred over Move. Merge, Agree, and Move are the three major operations in the computation under the minimalist assumptions. In more recent versions of the MP (e.g. Chomsky 2001, forthcoming), however, Move is renamed Internal Merge, and is argued to be freely available as well. If both Merge and Internal Merge are free, there seems to be no basis for Merge-over-Move Principle. Grammatical operations are subject to other universal principles, such as the **Locality Principle**, which requires operations to be local, and the **Earliest Principle**, which requires operations to apply as quickly as possible.

Now I take the subject-verb agreement (which is actually an agreement relation between T and DP) as an example to illustrate the agreement mechanisms just depicted. Chomsky assumes that T has uninterpretable ϕ features (represented as [*u*PER], [*u*NUM],[*u*GEN]) as well as an EPP feature; the goal DP has the matching interpretable ϕ features with specified values. Chomsky also assumes that the goal DP has a [*u*CASE] feature, which can only be valued by an ϕ -complete T. Accordingly, the initial derivational stages of sentences (3.3a) *A mouse was in the bath* and (3.3b) *There was a mouse in the bath* can be represented as below (examples from Chris Wilder’s presentation at *Chomsky Seminar*, 2006, NTNU):

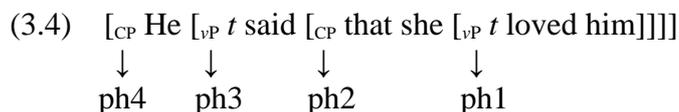


Note: ‘....’ means c-command.

In both cases, the DP *a mouse* enters the derivation with a full set of ϕ features and a [*u*CASE] feature; T has a full set of unvalued ϕ features and an EPP feature. T is therefore in a position to probe into its c-command domain for matching features. It finds the DP *a mouse*, and assumes the values of the valued features of the goal. The difference between sentences *a* and *b* resulted in the subsequent derivation stage, where the goal DP in *a* is forced to move to Spec TP in order to satisfy the EPP feature, whereas in *b*, the move operation does not occur; the EPP feature of T is satisfied by merging the expletive *there*. The two examples above therefore respectively illustrate a

case where agree yields displacement and a case where agreement is established in the base position, without involving movement.

Phase (ph) is also an important concept, introduced in the MP for the purpose of reducing computational complexity. Phases are propositions, including a verb phrase v^*P and a full clause CP. Accordingly, the sentence *He said that she loved him* constitutes four phases, as shown in (3.4):



Note: *t* stand for traces of (the closest) subject.

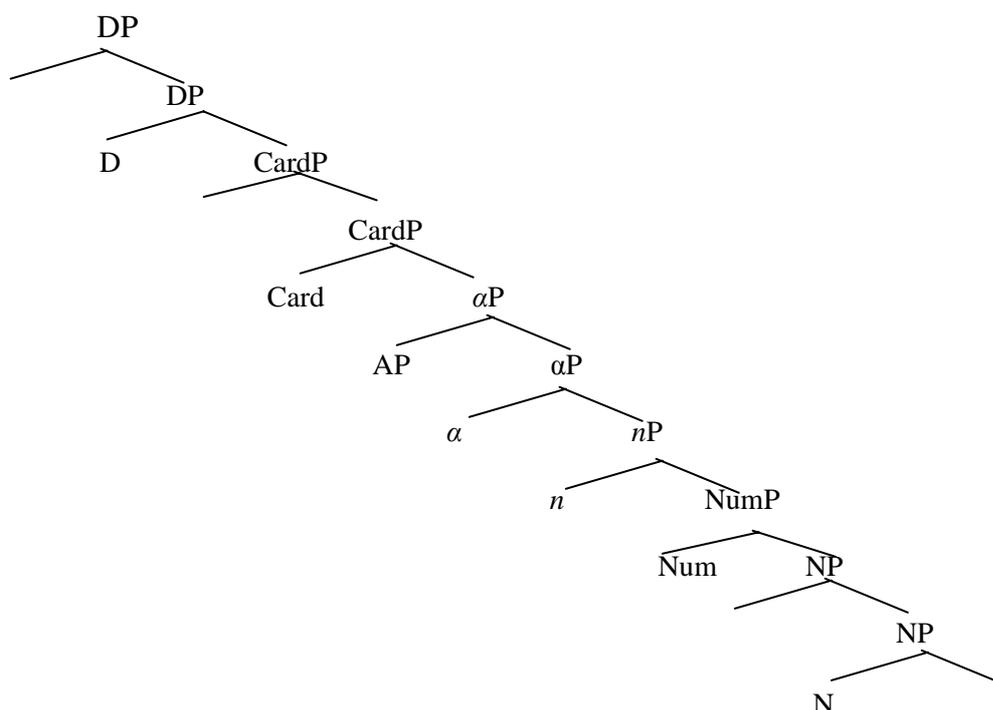
Derivations proceed phase by phase. Once a phase is completed, the constituents in the relevant domain are no longer visible to syntax, and not able to undergo any further syntactic operations. This constraint is referred to as Phase Impenetrability Condition (PIC). The operation Spell-out is cyclic, at the phase level. Spell-out removes uFs from the syntactic object and transfers it to the phonological component. By the end of a converging syntactic derivation, there will be no uFs left in narrow syntax.

3.3 Syntactic structure of Norwegian DPs

There have been many attempts to account for the syntactic structure of the Scandinavian DPs in general and Norwegian DPs in particular, especially the intriguing double definiteness construction (for some recent accounts, see for example, Delsing 1993, Kester 1996, Vangsnes 1999, Julien 2003, 2005). Among them, Julien (2003, 2005) provides the most thorough investigation of Scandinavian DPs. More importantly, her proposal is the most recent and the one within the minimalist framework. So the analysis of Norwegian DPs proposed in this work is based on Julien (2003, 2005). Following Julien, I assume that the (extended) Norwegian DPs have the syntactic structure shown in (3.5).

As we can see from (3.5), above NP is NumP. Its head Num, which signifies the number distinction, is the position where the suffixed number marker is generated. On top of NumP is *n*P, which is considered as the nominal counterpart of *v*P. So just as V obligatorily moves to *v*, N obligatorily head moves to *n*, via the Num head. Moreover, *n* hosts the suffixed definite article in Norwegian. Num and *n* heads presumably contain interface-interpretable features; in other words, the features enter into derivation with their values specified. For Num, the value is either singular or plural; for *n*, the value is either definite or indefinite. For expository purposes, I refer to them as [+pl] and [+def], respectively. [-pl] and [-def] are unmarked in Norwegian, and they have no phonological realization; [+pl] and [+def] are the marked ones, and they are realized as the suffixed number marker and the definite marker, respectively. In addition, there is a gender feature which is assumed to be inherent in the Norwegian nouns. Based on the two-way neuter and non-neuter (common) gender system in Norwegian, this gender feature is referred to as [+neut]. Thus when *n* completes a lower phase of the syntactic structure based on nominals, it carries a cluster of iFs, including [+pl], [+def], and [+neut].

(3.5)



When an adjective and a numeral is included, nP merges with an α head, which projects αP , with adjectival phrases (AP) sitting in its Spec. αP further merges with Card head, which projects CardP, with numerals and weak quantifiers sitting in its Spec. Finally on the topmost layer, we find DP layer, with D hosting the free pronominal determiners. DP, nP , NumP and NP layers are taken to be present in every Norwegian DP; αP and CardP, on the other hand, are generated only when they contain lexical material.

Under the assumption that syntactic features are morphologically expressed, Julien claims that the functional heads, namely D and α , contain uninterpretable gender, number, and definiteness features in Norwegian. These features enter into the derivation unvalued, I thus refer to them as [$uGEN$], [$uNUM$] and [$uDEF$] respectively, using [u] to represent unvalued features. At this juncture, some clarification is needed regarding the semantic interpretability of features. First and foremost, it is puzzling why grammatical gender on nominals is supposedly interpretable. It is well-known that the gender information on the noun is often not directly related to the real-world gender, hence makes no contribution to semantic interpretation. On the other hand, it appears to be counterintuitive to assume that D carries an [$uDEF$] feature, because it is commonly assumed that D head is a locus of an *interpretable* definiteness feature, with a value of either definite or indefinite. On the path to clear away the puzzles, it is useful to make a distinction between *inherent* features and *relational* features, and I agree with Yoon (1999), who notes that grammatical features that are inherent to a category are

interpretable; those that are relational to it are uninterpretable.¹³ Gender is inherent in the noun, thus grammatical gender on nominals is supposed to be interpretable. The gender feature on D, on the other hand, is an agreement category (as shown by D-N gender agreement), hence it is uninterpretable. The same reasoning can be applied to the interpretability of definiteness feature. Recall under Julien's account, an interpretable definiteness feature is encoded in the *n* head in Norwegian. Thus the definiteness feature on D, which exhibits definiteness agreement with *n*, is a relational feature in Norwegian and hence supposedly uninterpretable.

If interpretability of features can be understood this way, it follows that Danish D should not contain a [*u*DEF] feature on D, as Danish does not exhibit D-N definiteness agreement. Thus Danish D and Norwegian D are parameterized with respect to a [*u*DEF] feature: present in Norwegian; absent in Danish. Readers who have already been familiar with Norwegian DP structures will immediately ask the question: How can we account for the fact the Norwegian DPs in the formal style may have the same form as Danish DPs? I retain the claim that the syntactic representations of DPs in the two languages are different, and argue instead that the difference between Norwegian definite DPs in the formal style and Danish definite DPs actually reflects a distinction between *null* morpheme and *absence* of morpheme (cf. Lardiere 2000). In Norwegian, D contains a [*u*DEF] feature, and *n* has a matching [+DEF] feature, so a Probe-Goal relation will be established between D and N in order to delete the unvalued feature on D. In normal cases, the [+DEF] feature is realized as an overt suffix on the noun, whilst in the formal style, the feature may have no phonological realization; in other words, the suffix is realized as a null morpheme. In Danish, by contrast, the [*u*DEF] feature is not specified on D at all, and *n* never has a phonological realization. That is to say, the suffix is totally absent in Danish.

This being clear, we now go on to tackle another problem: how can the seemingly defective inflection *-e* on attributive adjectives indicate that the α head in Norwegian is fully specified with [*u*GEN], [*u*NUM] and [*u*DEF] features, given the assumption that uFs are morphologically expressed? Julien explains that although there are many different feature combinations (out of the three features: gender, number and definiteness) in Norwegian, there are only three possible realizations on adjective: *-t*, zero (*-∅*), or *-e* (cf. Chapter 2). *-t* marks neuter singular indefinite; zero (*-∅*) marks common singular indefinite; *-e* is an unspecified realization of agreement. Invoking the concept *impoverishment* of Distributed Morphology (cf. Halle 1997), Julien argues that in Norwegian, SINGULAR features are deleted in the context of DEFINITE, which makes *-t* and zero unavailable, and *-e* appears instead. In this view, *-e* can be fully specified for number, gender and definiteness features. The weak inflection *-e* is just the default realization of this feature cluster.

If this line of argument is on the right track, head D and head α are Probes, seeking for the closest Goal they can agree with. This agreement does not necessarily require overt movement. Recall that under minimalist assumptions, long distance agreement is possible. Whether or not Move is involved in the derivation of DPs is mainly dependent

¹³ An apparent point of departure from this view is the number feature on nominals. It is generally held that number is not an inherent property of nouns. But Chomsky (1999) notes that the number feature on nominals, along with other ϕ features, are presumably interpretable. I think this issue can be explained this way: the number feature is an inherent property of the Num head. When the N head moves to Num, it picks up the number feature on Num and thereby gets either a singular or a plural interpretation.

on the following two conditions: 1) whether D offers a position for movement, and 2) what kind of category D seeks. The first condition is satisfied in Norwegian. Julien proposes that the reference of DP in Norwegian depends on D, and because of this, D (or Spec DP) must have a phonological realization in order for the DP to be referential. This means Spec DP can be a potential landing site for movement operations.

As to what category D seeks, Julien argues in line with Baker (2003), that DP must have a nominal category feature, which is understood as the ability to have a referential index. She further assumes that *n*P has the required nominal category feature when it has a phonological realization, i.e. when the definiteness feature is spelt out. Accordingly, it is possible that *n*P be attracted to Spec-DP. In building minimal definite DPs, i.e. definite DPs without prenominal modifiers, the operation Agree applies as soon as D is merged above *n*P, and eliminates the uFs on D. At this juncture, there are two alternatives through which the referential requirement of D can be satisfied. One is to move *n*P to Spec DP; the other is to spell out D itself. Julien argues that moving *n*P to Spec DP is preferred to spelling out D itself, because internal merge is less costly than introducing an extra lexical item. As a result, minimal definite DPs end up with a suffixed determiner, rather than a prenominal determiner.

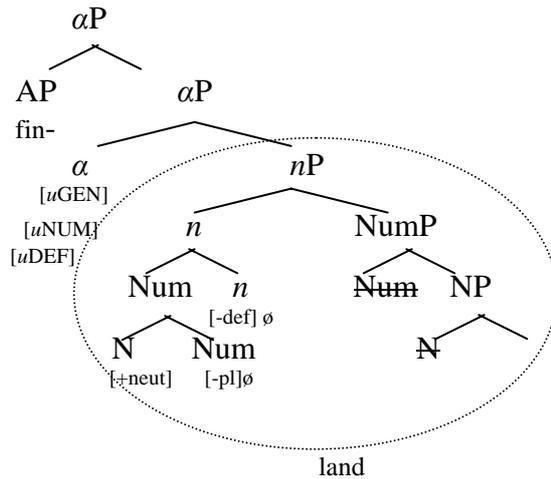
When a definite DP is modified (by an attributive adjective, for example), however, the α head is merged above *n*P, and a Probe-Goal relation is first established between α and *n*P, and via Agree the uFs of the Probe α would be valued and rendered interpretable. Presumably, AP shares all the features with the α head. So when D is merged above α P, which has an AP in its Spec, the AP will become the closest Goal for the Probe D. Thus movement of *n*P to Spec-DP is blocked. Meanwhile moving AP to Spec DP cannot satisfy the requirement of D, because an AP in Norwegian does not have a nominal category feature. The only solution then is to spell out D itself as a prenominal determiner. This derivation leads to double marking of definiteness, giving rise to the double definiteness construction. A similar Probe-Goal relation is established in singular indefinite DPs. But since *n* has no phonological realization in indefinite DPs, *n*P does not have a nominal category feature, and hence is not entitled to specify the reference of D. It then becomes necessary to spell out D, whether or not an adjective is present. Because of the matching of features in the process of derivation, the functional head D shares all the features with α , which in turn has features identical to the *n*P complex.

As we see, different operations are involved in the derivation of indefinite DPs, definite DPs without a prenominal modifier, and definite DPs with prenominal modifiers. I thus will illustrate the three types of DPs each with a tree diagram. Tree diagram (3.6) illustrates the major steps involved in the derivation of the indefinite phrase *et fint land* ‘a nice country’, (3.7) illustrates the major derivational steps of a minimal definite phrase *landet* ‘the country’, and (3.8) illustrates the major derivational steps of *det fine landet* ‘the nice country’, a DP of ‘double definiteness’ construction. Only the syntactically relevant features are shown. In order to simplify exposition, here I only include syntactic representation after the successive movement of N to *n* has been completed (indicated by a circle with dotted lines). Also note that since the Earliness Principle specifies that operations apply as early as possible in a derivation, it follows

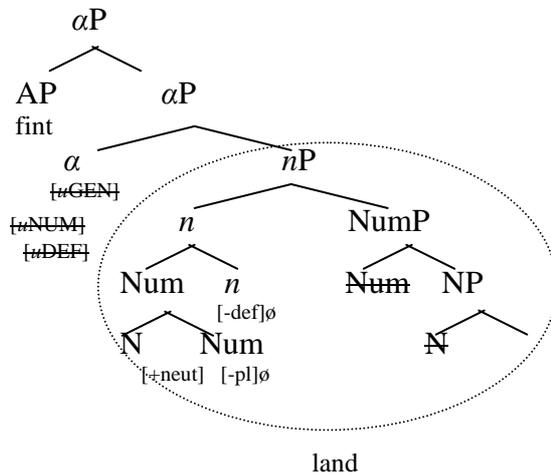
that Agree applies as soon as the Probe is merged. I divided the operations into individual steps just for expository purposes.¹⁴

(3.6) *et fint land* ‘a nice country’

- a. MERGE α and nP to form an αP , with AP sitting in the SPEC of αP . Head α carries [$uGEN$], [$uNUM$], and [$uDEF$] features. The nP complex carries an interpretable feature bundle [$+neut$ $-pl$ $-def$], which spells out a null morpheme.

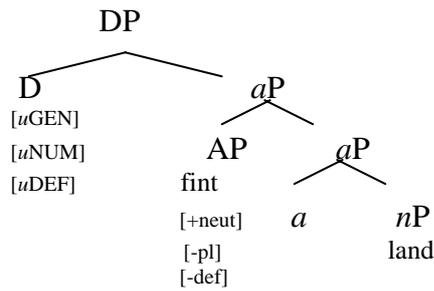


- b. AGREE { α =Probe, nP complex=Goal, valuation/deletion of [$uGEN$],[$uNUM$], and [$uDEF$] on α }. Now feature values of nP have been copied on to the head α , so that α carries an interpretable feature bundle [$+neut$ $-pl$ $-def$]. AP agrees with α , and the feature bundle on AP spells out as $-t$.

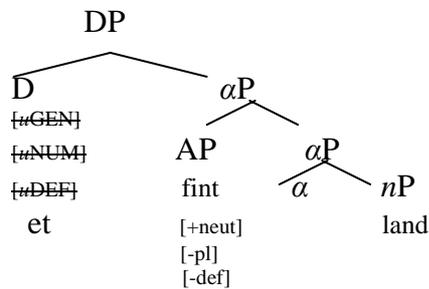


- c. MERGE D and αP . Head D carries [$uGEN$],[$uNUM$], and [$uDEF$] features; AP carries an interpretable feature bundle [$+neut$, $-pl$, $-def$].

¹⁴ A symbol ‘ \emptyset ’ is used to indicate that the related feature has no phonological realization. A strikethrough on the head indicates that the head has undergone head movement to another head, and left a copy in its original position. A strikethrough on the uFs indicates that a given feature has been valued and marked for deletion at LF.



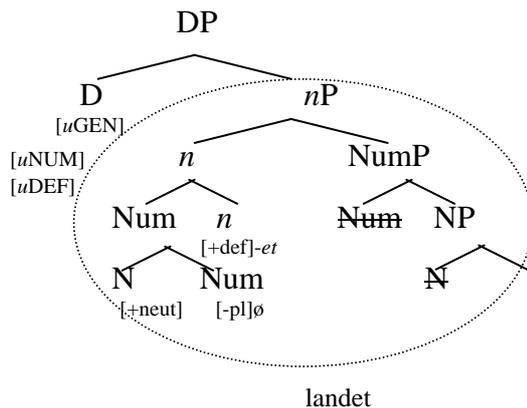
- d. AGREE {D=Probe, AP=Goal, valuation/deletion of [uGEN],[uNUM] and [uDEF] on D}. Feature values of AP have been copied on to the head D, so that D carries an interpretable feature bundle [+neut, -pl, -def]. As AP presumably does not have a nominal category feature, it cannot be attracted to Spec DP. Instead D itself is spelt out. These feature values on D spell out as *et*.



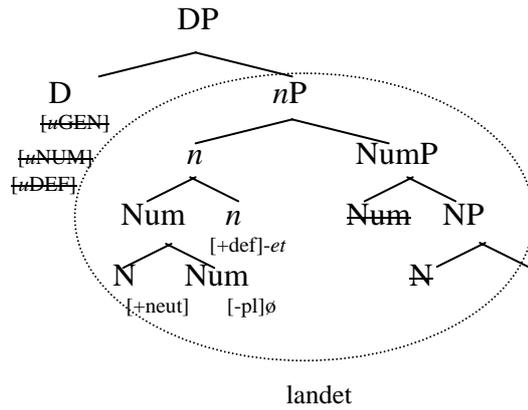
- e. TRANSFR...end of derivation

(3.7) *landet* 'the country'

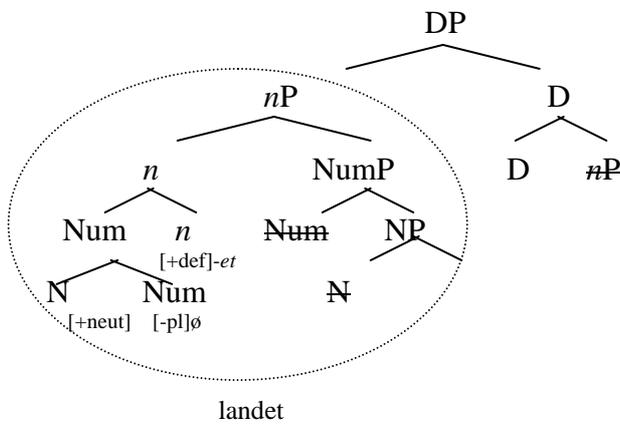
- a. MERGE D and nP. Head D carries [uGEN],[uNUM], and [uDEF] features; the nP complex carries an interpretable feature bundle [+neut -pl +def], which spells out an *-et* morpheme.



- b. AGREE {D=Probe, nP complex=Goal, valuation/deletion of [uGEN],[uNUM] and [uDEF] on D}. Now feature values of nP have been copied on to the head D, so that D carries an interpretable feature bundle [+neut -pl +def].



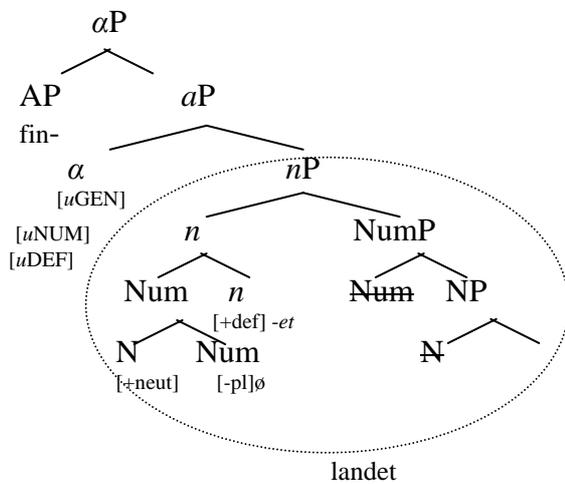
c. MOVE *nP* to Spec-DP to satisfy the referential requirement of *D*.



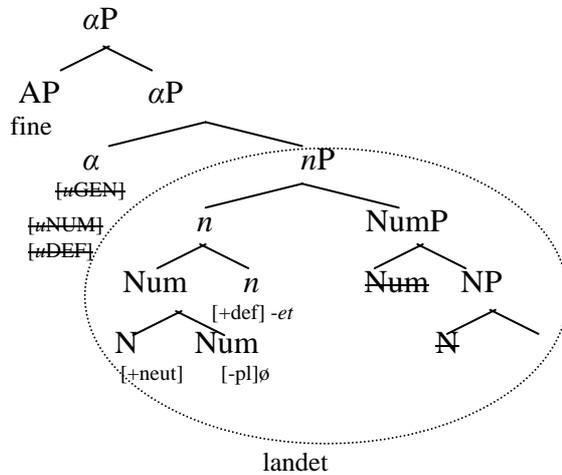
d. TRANSFR...end of derivation

(3.8) *det fine landet* ‘the nice country’

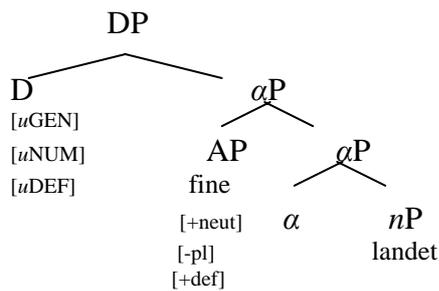
a. MERGE α and *nP* to form an αP , with AP sitting in the SPEC of αP . Head α carries [*u*GEN], [*u*NUM] and [*u*DEF] features. The *nP* complex carries an interpretable feature bundle [+neut -pl +def], which spells out as *-et*.



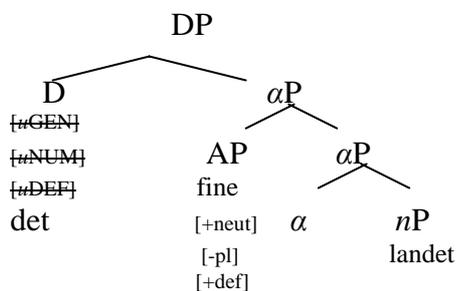
- b. AGREE { α =Probe, nP complex=Goal, valuation/deletion of [$uGEN$],[$uNUM$], and [$uDEF$] on a }. Feature values of nP have been copied on to the head α , so that α carries an interpretable feature bundle [+neut -pl +def]. AP agrees with α , and the feature bundle on AP spells out as $-e$.



- c. MERGE D and αP . Head D carries [$uGEN$],[$uNUM$] and [$uDEF$] features; AP carries an interpretable feature bundle [+neut -pl +def].



- d. AGREE [D=Probe, AP=Goal, valuation/deletion of [$uGEN$],[$uNUM$], and [$uDEF$] on D]. Feature values of AP have been copied on to the head D, so that D carries an interpretable feature bundle [+neut -pl +def]. As AP does not have a nominal category feature, it cannot be attracted to Spec DP. Instead D itself is spelt out. These feature values on D spell out as det .



- e. TRANSFR...end of derivation

3.4 Syntactic structure of Italian/Spanish DPs

Julien (2005), following Chomsky, assumes that the basic syntactic structure of nominal phrases is uniform across languages. Variation, she claims, has to do with the feature makeup of the individual functional heads, with what is spelt out and where, and with the movement operations that may or may not take place inside the nominal phrases. Following this line of argument, I assume that the DP structures of Italian/Spanish are fundamentally the same with that of Norwegian, with variation only in feature makeup of functional heads D , α , and n . The variation is actually due to one parametric difference, i.e. the presence of [μ DEF] in Norwegian and the absence of it in Italian/Spanish. As we will see, this parametric difference has some consequences on the syntactic operations involved in the derivation of Italian/Spanish DPs.

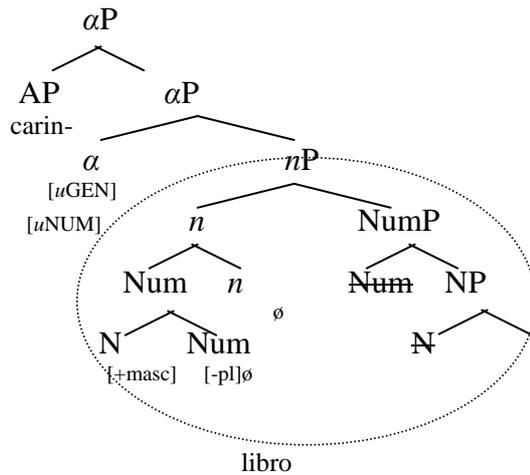
In Italian/Spanish there is no definiteness concord between D and n or between α and n . This is taken as an indication that n never gets a phonological realization. Hence it does not encode the definiteness information. Rather, D is the locus of the interpretable definiteness (i.e. [\pm def]) feature. Moreover, D and α carry uninterpretable number and gender features, which must be valued and deleted by Agree. Crucially this operation does not induce Move, but rather is established on the base positions of α , nP , and D . Recall that the constituents on D must have a nominal category feature in order to identify the reference of D . Since n never has a phonological realization, nP does not have a nominal category feature. As a result, nP complex cannot be attracted to Spec-DP. The solution therefore, is to spell out D itself. Thus we see that in the derivation of Italian/Spanish DPs, Merge and Agree apply, but no operation Move is involved except for the obligatory movement from N to n via Num¹⁵. Because of this, the presence of an adjective or a numeral in the Italian/Spanish DPs has no consequences in the syntactic operations involved.

To illustrate, I take an Italian DP *il libro carino* ‘the nice book’ as an example. Italian/Spanish nouns distinguish masculine and feminine gender, so the gender feature on nouns is referred to as [\pm masc] in the exposition.

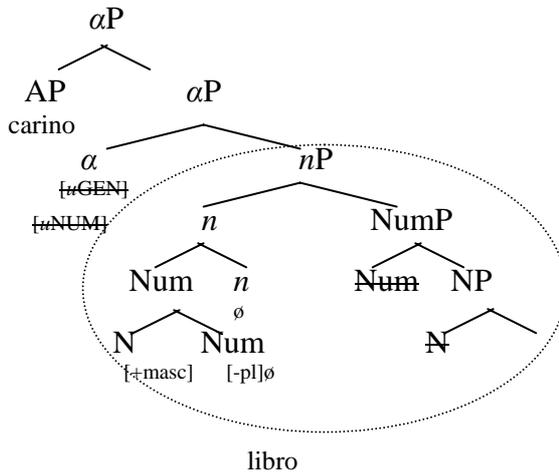
(3.9) Italian: *il libro carino* ‘the nice book’

- a. MERGE α and nP to form an αP , with AP sitting in the SPEC of αP . Head α carries [μ GEN] and [μ NUM] features. The nP complex carries an interpretable feature bundle [$+$ masc $-$ pl], which spells out a null morpheme.

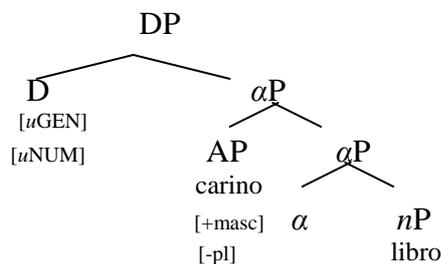
¹⁵ But note that in Italian/Spanish, attributive adjectives can appear in the postnominal positions. Presumably, this is an indication that the noun further moves upwards in the DP. But this variation does not affect the agreement pattern under discussion, and therefore is not of concern here.



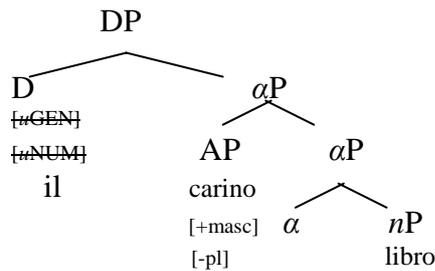
- b. AGREE { α =Probe, nP complex=Goal, valuation/deletion of [uGEN] and [uNUM] on α }. Feature values of nP have been copied on to the head α , so that α carries an interpretable feature bundle [+masc, -pl]. AP agrees with α , and the feature bundle on AP spells out as *-o*.



- c. MERGE D and αP . Head D carries [uGEN] and [uNUM] features; AP carries an interpretable feature bundle [+masc -pl].



- d. AGREE {D=Probe, AP=Goal, valuation/deletion of [uGEN] and [uNUM] on D}. Feature values of AP have been copied on to the head D, so that D carries an interpretable feature bundle [+masc -pl]. As AP does not have a nominal category feature, it cannot be attracted to Spec DP. Instead D itself is spelled out. These feature values on D spell out as *il*.



e. TRANSFR...end of derivation

3.5 Syntactic structure of English DPs

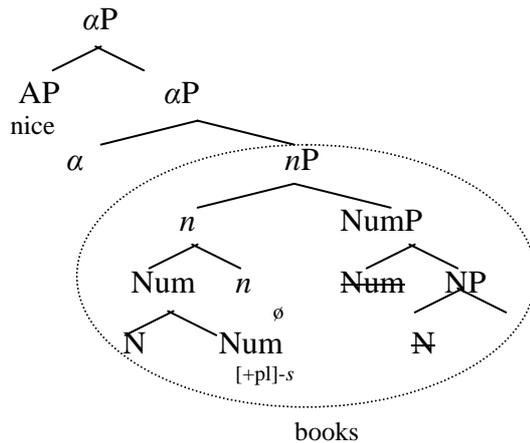
From the above section we have seen that English does not have gender or definiteness concord between determiners and nouns. Moreover, English adjectives do not inflect at all. But English does exhibit a number agreement between nouns and demonstratives. These observations reflect a number of parametric differences between English and Norwegian DPs. First, the α head in English does not carry any uFs; second, D, but not n , is the locus of interpretable definiteness feature in English. Finally, the only uninterpretable feature is [u NUM] on D; [u DEF] and [u GEN] features are not present in English. To relate these grammatical properties of English with Julien's syntactic model, we see that neither n or α has phonological realization in English. As the case in Italian/Spanish, n does not have a phonological realization, so the n P will not suffice to identify D, and hence it can not move to Spec-DP. As a result, D itself is spelt out whether or not an adjective or numeral is present. Furthermore, as α head does not carry any unvalued features, it remains inactive, and is "frozen in place" (Chomsky 1999:6). Thus the operation Agree does not apply at this point of the derivation. Only when D is merged in English, the [u NUM] feature it carries renders it active, and the Probe-Goal relation between D and n P will be established. Via Agree, this [u NUM] feature in D is valued and eliminated. As we see, compared with the steps in the derivation of Norwegian DPs, the derivation of English DPs involves fewer Agree operations, and does not invoke Move at all.

Again I use tree diagrams to illustrate the major steps in the derivation of English DPs. This time I take *these nice books* as an example.¹⁶

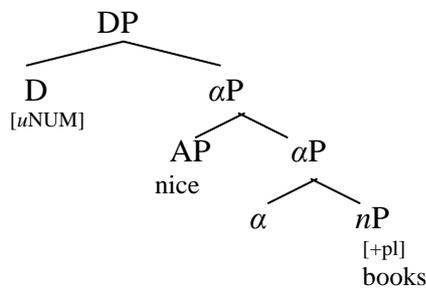
(3.10) English: *these nice books*

- a. MERGE α and n P to form an α P, with AP sitting in the SPEC of α P. Head α does not carry any uFs, so operation Agree does not apply at this conjuncture.

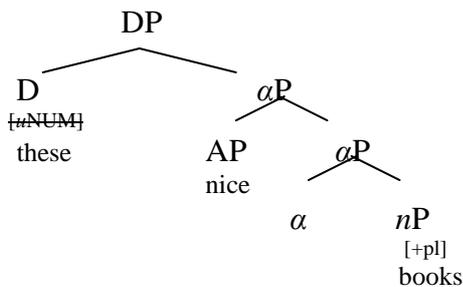
¹⁶ Julien (2005) argues that demonstratives have an individual project above DP. She refers to it as DemP. To simplify exposition, I do not make such a distinction. Thus in my analysis D is more inclusive than in Julien (2005).



- b. MERGE D and αP . Head D carries a $[uNUM]$ feature; nP carries an interpretable feature $[+pl]$



- c. AGREE {D=Probe, nP =Goal, valuation/deletion of $[uNUM]$ on D}. The feature value of nP has been copied on to the head D, so that D carries an interpretable feature $[+pl]$. As nP does not have a nominal category feature, it cannot be attracted to Spec DP. Instead D itself is spelt out. The feature value on D spells out as *these*.



- d. TRANSFR...end of derivation

3.6 A note on Chinese DPs

Turning to Mandarin Chinese, we find that Chinese nominal phrases do not readily fit into Julien's model. As a non-inflectional language, Chinese does not have any definiteness marker, nor has it any gender distinction on its nominals. Arguably, it does not have a number marker either. Instead, a classifier has to be used with nouns in

counting. In the previous chapter, I argued in line with Yang (2004) that the suffix *-men* is a suffixal classifier, rather than a plural marker, and that D in Chinese does not contain a number feature. Thus there is no agreement relation between D and N in terms of number in Chinese. As a result, Agree does not apply in the derivation of Chinese DPs.

Taking all these facts about Chinese together, it seems that the functional projections proposed in Julien's model, including DP, α P, *n*P, and NumP, are not morphologically supported in Chinese. This great variation in morphological properties between Chinese and Indo-European languages has led many researchers to assume that there exists no universal syntactic structure across languages. Indeed many linguists argue that the maximal projection of Chinese nominal phrases is not a DP. Chen and Sybesma (1999), for example, propose that Chinese nominal phrases are maximally NumeralPs, which select a ClassifierP.¹⁷ Huang *et al.* (to appear), however, assumes that all languages have identical nominal structures, and argue convincingly that the maximal projection of Chinese nominal phrases is a DP.

In this study, I follow the assumption that the basic syntactic structure is invariant and adopt Julien's DP structure for Chinese nominal phrases. To incorporate the facts about Chinese into her model, Julien speculates that the classifier could be a realization of the Card head, which has a numeral phrase in its Spec, or it could belong to the numeral phrase itself. Clearly, it requires some further elaboration for Julien's argument to be convincing. Particularly, how could Julien's model accommodate the co-occurrence constraints of *-men* and *xie* with numerals? Consideration of this, however, would take us too far afield. For present purposes, suffice it to point out that no functional heads in Chinese DPs carry any uninterpretable features, and that Agree does not apply in the derivation of Chinese DPs.

3.7 Summary and a road ahead

In this chapter, I have presented an analysis of the DP structures of Norwegian, Italian/Spanish, English, and Chinese based on Julien's (2003, 2005) model on nominal phrases in Scandinavian. Julien argues that the functional heads D and α contain uninterpretable gender, number and definiteness features in Norwegian. These features enter into the derivation unvalued, and must be valued via Agree with an element that has valued features. She also assumes that the basic syntactic structure of nominal phrases is uniform across languages, with variation only in feature makeup of functional heads D, α , and *n*. Following this line of argument, we assume that the DP structures of Italian/Spanish, English and Chinese are fundamentally the same as that of Norwegian. Variations among these languages are attributed to presence or absence of uFs on D and α heads. We have shown that the parameterized features, especially parameterized uFs, give rise to variations in the syntactic operations involved in derivations.

The parameterized uninterpretable features in target and source languages are summarized in Table 3.1, based on the facts whether or not the languages have overt gender, number and definiteness agreement between determiners, adjectives, and nouns.

¹⁷ Their NumeralP is equivalent of Julien's CardP.

Table 3.1. Parameterized uFs in the target and the source languages

	D			α		
	[<i>u</i> NUM]	[<i>u</i> GEN]	[<i>u</i> DEF]	[<i>u</i> NUM]	[<i>u</i> GEN]	[<i>u</i> DEF]
Norwegian	+	+	+	+	+	+
Italian/Spanish	+	+	-	+	+	-
English	+	-	-	-	-	-
Chinese	-	-	-	-	-	-

Key: + Present in language; - absent in language

As we can see from Table 3.1, the target language, Norwegian, has [*u*NUM], [*u*GEN], and [*u*DEF] features both on the D and the α heads; Italian/Spanish has [*u*NUM] and [*u*GEN] on D and α , but not [*u*DEF] feature either on D or on α ; The only uninterpretable feature English has is [*u*NUM] feature on D; Chinese has none of the uFs either on the D or on the α head.

In this study, I ask the question whether or not interlanguage grammars of endstate L2 Norwegian speakers have the same properties as the native-speaker grammars. In particular, will the [*u*NUM], [*u*GEN], and [*u*DEF] features in the target language be present in L2 speakers' interlanguage grammar irrespective of status of the corresponding feature(s) in their L1s? In addition, I will also investigate whether L2 learners would be sensitive to DP internal agreement violations when processing L2 Norwegian, and whether their sensitivity to a certain agreement cue would be dependent on the presence of the formal feature in their L1s. Given the parametric differences between the target and the source languages, I am in a position to examine the fundamental issues in the ultimate attainment of second languages acquisition: L1 transfer effects and L2 speakers' competence at the level of grammatical knowledge as well as processing.

4 *Literature review*

4.1 Introduction

For over a decade L2 researchers doing ultimate attainment studies have been debating whether parameterized linguistic properties are available to adult L2 learners or whether they are subject to a critical period. The present chapter reviews two competing models in the current SLA literature, namely the *Failed Functional Features Hypothesis* (*FFFH*, e.g. Tsimpli & Roussou 1991; Hawkins and Chan 1997; Hawkins 2001b, 2004; Franceschina 2001; Hawkins and Liszka 2003; Tsimpli 2003), and the *Full Transfer Full Access Hypothesis* (*FTFA*, Schwartz and Sprouse 1996; White 2003). As we will see, the two positions often carry contrasting claims about what linguistic properties can or cannot be eventually acquired. As such, they are often found in opposing camps, such as *no parameter-resetting* vs. *parameter resetting* (White 2000, 2003), and more recently, *features missing despite appearances* vs. *features present despite appearances* (Hawkins 2004).

Originally formulated under the Government and Binding (GB) framework, the *FFFH* claims that L2 learners' deficits are representational, and that UG-based parameters, particularly parameterized functional features, which are not realized in L1s, cannot be reset to the L2 settings after a critical period; the *FTFA*, on the other hand, argues that L1 grammar constitutes the L2 initial state (hence the term *full transfer*), but it allows for the possibility of UG-based grammar restructuring (i.e. parameter resetting) in response to L2 input. It predicts that eventually nativelike L2 interlanguage grammatical representations are attainable. As general syntactic theory gradually developed from the GB theory to the Minimalist Program (MP), the two models have accordingly made more fine-grained claims as to what aspects of the target language are subject to maturational constraints. The *FFFH* makes a crucial distinction between interpretable and uninterpretable features, and claims it is the parameterized uninterpretable features, not the interpretable counterparts, that are affected by a critical period (hence *FFFH* is also referred to as the *Interpretability Hypothesis*, see Tsimpli 2003). The *FTFA* maintains that the whole set of formal features provided by UG, interpretable and uninterpretable, continue to be available to adult L2 learners. In order to account for the observed NS/NNS divergence, researchers espousing the *FTFA* put forward two proposals, namely, the *Missing Surface Inflection Hypothesis* (*MSIH*, Lardiere 1998a, 1998b, 2006a; Prévost and White 2000) and the *Prosodic Transfer*

Hypothesis (PTH, Goad *et al.* 2003; Goad and White 2004). The MSIH argues for a disassociation between syntax and morphology/phonology, and maintains that it is the interface with the peripheral morpho/phonological components that is susceptible to maturational constraints. The PTH claims the ultimate attainment of L2 speakers is constrained by L1 prosodic representations. I will review and assess the studies where the FFFH and the FTFA were originally formulated, and some of the influential studies contributing to the refinement/development of the two models.

SLA theories will not be complete without a theory on L2 processing. The processing theory which is of interest to the current study is the *Shallow Structure Hypothesis* (SSH, Clahsen and Felser 2006a, 2006b). According to the SSH, L2 learners have recourse to different processing routes as compared to native speakers. Whereas L1 processing typically involves full syntactic representations, adult L2 processing is restricted to shallow computation, relying on lexical/pragmatic information. But crucially this L1/L2 processing difference is claimed to be restricted to non-local domains. For local domain properties, nativelike L2 processing is predicted to be possible.

4.2 Failed Functional Features Hypothesis (FFFH)

4.2.1 Hawkins and Chan (1997)

We begin with an examination of a study by Hawkins and Chan (1997), where the failed functional features hypothesis (FFFH) was initially proposed. The FFFH, arguing in line with the no-parameter resetting hypothesis (Tsimplici and Roussou 1991), claims specifically that grammar building in second language acquisition is UG constrained, but *parametric functional features* resist resetting. This hypothesis is tested using two groups of L2 learners of English (L1 Cantonese¹⁸ and L1 French). The aspect of language under investigation is English restrictive relative clauses (RRCs). Adopting the GB framework, Hawkins and Chan assume that English/French and Cantonese relative Cs all have a predicative feature (i.e. C [pred]), but are parameterized with respect to a [wh] feature: the feature is present in English and French, but absent in Cantonese. This parametric variation gives rise to a number of characteristic differences between English and Cantonese RRCs. First, it is assumed that English RRCs involve the movement of a *wh*-phrase (also called a *wh*-operator) to the specifier position of CP in the embedded clause. This movement leaves a trace in the position from which the *wh*-phrase has moved (cf. example 4.1a). Since the moved *wh*-phrase is a trace, a resumptive pronoun in this position is disallowed, thus the ungrammaticality of sentence (4.1b). Moreover, English RRCs disallow doubly filled C (cf. example 4.1c), presumably due to a feature clash between a *wh*-phrase (with a + [wh] feature) and the complementizer *that* (with a - [wh] feature).

¹⁸ Cantonese is one of the Chinese dialects, mainly spoken in Guangzhou province and in Hong Kong. As far as the relative clause is concerned, Hawkins and Chan (1997) assume that the grammar of Cantonese does not differ much from that of Mandarin, i.e. standard Chinese.

- (4.1) a. The girl [_{CP} who_i [_{IP}I like t_i]] is here. (Example from Hawkins and Chan, 1997)
 b. *The girl [_{CP} who_i [_{IP}I like her_i]] is here. (Resumptive pronoun)
 c. * The girl [_{CP} who_i that [_{IP}I like t_i]] is here. (Doubly filled C)

In addition, *wh*-operator movement in English is subject to subjacency constraints (or ‘shortest move’). Moving a *wh*-phrase from a *wh*-island or a complex NP results in subjacency violations, as shown in (4.2 *a* & *b*) below:

- (4.2) a. * The man who Mary told me [when she will visit <who>] is here.
 (*Wh*-island)
 b.* The boy who Mary described [the way [Bill attacked <who>]] is here.
 (Complex NP)

By contrast, Chinese RRCs do not involve a *wh*-phrase. A complementizer (*de* in Mandarin; *ge* in Cantonese) immediately precedes the modified DP. In addition, resumptive pronouns are optional. Following Huang (1984, 1989) and Xu & Langendoen (1985), Hawkins and Chan assume that Chinese RRCs involve the base-generation of a null topic that binds a null or resumptive pronoun, rather than movement. This is illustrated in (4.3):

- (4.3) [_{CP} Ø-Top_i [_{IP} wo xihuan pro_i/ta_i] de] nei ge nühai_i
 Null topic I like e/her C the CL girl
 ‘The girl that I like.’

Hawkins and Chan tested Cantonese-speaking L2 learners of English at three proficiency levels (elementary, intermediate, advanced) and compared them with age- and proficiency-matched French-speaking L2 learners and native controls. The task they use is a grammaticality judgment task, involving 20 grammatical and 39 ungrammatical RRCs (including *wh*-island violations, complex NP violations, resumptive pronouns and doubly-filled Cs). The results of the Chinese groups are presented in Table 4.1 and 4.2; the results of French groups in table 4.3 and 4.4.

Table 4.1. Chinese group’s accuracy in judging English RRCs (%)

	Gramm RRCs	*Doubly-filled C	*Resumptive	Prons
Elemen. Chinese (n=47)	56	50	38	
Intermed. Chinese (n=46)	67	68	55	
Advanced Chinese (n=54)	79	83	90	
NS controls (n=32)	96	99	98	

Table 4.2. Chinese group’s accuracy in rejecting *wh*-island and complex NP violations

	* <i>wh</i> -island	*complexNP
Elemen. Chinese (n=47)	63	71
Intermed. Chinese (n=46)	54	61
Advanced Chinese (n=54)	41	38
NS controls (n=32)	98	85

Table 4.3. French group's accuracy in judging English RRCs (%)

	Gramm RRCs	*Doubly-filled C	*ResumptiveProns
Elementary Fren (n=47)	81	91	81
Intermed. Fren (n=46)	88	95	90
Advanced Fren (n=54)	92	98	96
NS controls (n=32)	96	99	98

Table 4.4. French group's accuracy in rejecting *wh*-island and complex NP violations

	* <i>wh</i> -island	*complexNP
Elementary Fren (n=47)	59	72
Intermed. Fren (n=46)	66	79
Advanced Fren (n=54)	85	90
NS controls (n=32)	98	85

Table 4.1 suggests that Cantonese speakers of English appear, as proficiency level increases, to approximate to the target language, accepting grammatical RRCs and rejecting doubly filled Cs and resumptive pronouns. However, Table 4.2 shows that Cantonese speakers appear to become less target-like with proficiency in detecting the ungrammaticality of subjacency violations. This is in contrast to French-speaking learners of English whose accuracy on and ability to correct Subjacency violations increase with proficiency (see Table 4.3 and 4.4). Why should such a contrast exist? Hawkins and Chan claim that adult L2 learners are constrained by the feature specifications of functional categories in their L1s. French, like English, allows *wh*-operator movement; whereas Chinese does not. So French-learners can achieve target-like grammatical representations. Chinese learners, by contrast, are confined to an interlanguage that is L1-like, that is, topic-antecedent-bound pronoun structures, rather than operator-bound variable structure. The compelling evidence of Chinese-like representations comes from the performance of the advanced learners, who did not show sensitivity to violations of movement constraints (i.e. subjacency violations), despite their apparent target-like performance on other properties. Based on these findings, Hawkins and Chan conclude that features (including feature strength) of functional categories that determine parametric options are subject to a critical period.

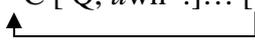
A critique Hawkins and Chan's study has received is that they did not offer an explanation for the apparent nativelike performance of the Cantonese learners as suggested by results from 4.1 above (see White 2003:125; Lardiere 2006:162, for relevant comments). At the methodological level, a possible drawback of Hawkins and Chan's study is that they did not test end-state Cantonese-speaking learners' intuitions on English RRCs. Though the advanced Cantonese group in their study had an exposure to English for an average of 14.6 years, their exposure to English was restricted to classroom settings, and therefore cannot be viewed as a reliable predictor for ultimate attainment, as Hawkins and Chan themselves argued. Without data from end-state L2 learners, any discussion about ultimate attainment is bound to be speculative in nature. Another weakness, later pointed out by Hawkins (2004) himself, is that the description 'features that determine parametric differences between languages' is vague, and needs to be defined more carefully. As general syntactic theory has developed, Hawkins

(2004) has re-examined the study and advanced a new version of FFFH in minimalist terms. We will turn to the updated version of FFFH in the next section.

4.2.2 Hawkins (2004); Hawkins and Hattori (2006)

Hawkins (2004) proposes a modified version of the FFFH in the spirit of the most recent development of generative linguistic theory, the Minimalist Program. Adopting Adger's (2003) account of *wh*-movement in English interrogatives, Hawkins assumes that C[Q] in English has a strong (represented with an asterisk) uninterpretable *wh*-([*uwh**]) feature, which drives the movement of a *wh*-phrase (with an interpretable [*wh*] feature) to Spec CP through an Agree operation. By extension, C[pred] in English RRCs also has a [*uwh**] feature. With an agreement mechanism similar to interrogatives, *wh*-phrases in RRCs move to Spec CP of the embedded clause. By contrast, no [*uwh*] feature is associated with C[pred] in Cantonese; topic-antecedent-bound pronoun structures are assumed for Cantonese RRCs. Under these syntactic assumptions, Hawkins reinterpreted the L2 data in Hawkins and Chan (1997), claiming that the problem for the Cantonese speakers of English is to establish a strong [*uwh**] feature associated with English C[pred]. On this basis, he put forward a new version of the FFFH, according to which only *uninterpretable syntactic features* (and feature strength) that have not been selected during first language acquisition, will not be available for L2 grammar construction; all other aspects of UG, including the computational devices, their associated operating principles, interpretable syntactic features and uninterpretable features selected into L1s, remain available in subsequent language acquisition. The modified version of the FFFH is also called the *Interpretability Hypothesis* (IH, see also Tsimpli 2003, Hawkins and Hattori 2006). Implicated in this hypothesis is the view that adult L2 learners' deficits are representational. In other words, after a critical period it is impossible for the L2 grammar to represent uninterpretable features not selected into the L1. Hence, IH is also referred to as the *Representational Deficit Hypothesis* (RDH, e.g. Hawkins 2003; Hawkins and Liszka 2003).

Hawkins (2004) provides further evidence supporting the IH, citing a study by Hawkins and Hattori (in progress; later, Hawkins and Hattori 2006), which investigates the acquisition of English multiple *wh*-questions by high proficiency Japanese-speaking learners of English (JSE). Hawkins and Hattori assume that English, a language that exhibits *wh*-movement, and Japanese, a *wh-in-situ* language, are parameterized with respect to feature strength of a [*uwh*] feature associated with C [Q]: strong in English, weak in Japanese. In addition to subadjacency effects (see the previous section), English multiple *wh*-questions show superiority effects, that is, C in multiple *wh*-questions attracts movement of the *closest wh*-expression which it c-commands, as a consequence of the *Attract Closest Principle*. This is schematically represented in (4.4a). A superiority violation is illustrated in (4.4b).

- (4.4) a. C [Q, *uwh**:]... [*wh*]1 ... [*wh*]2

 b. * What did John say [who ate <what>]?

In accordance with the IH, Hawkins and Hattori predict that JSE will not have access to the strength part of the [*uwh**] feature in constructing their ILGs, so they will not be sensitive to subadjacency and superiority violations in English. This prediction is tested on

19 JSEs and a control group of 19 native speakers of English (NSE), using a truth value judgment task, which took the form story + biclausal question + 3 answers (see below for an example).

Sophie was angry. Her holiday had been ruined because the hotel she had booked through a travel agency was full, and she had to sleep in a tent. Sophie's brother was a friend of Norman who owned the travel agency. He spoke to Norman on Thursday and told him that Sophie would be phoning his manager, Mrs Smith, the following day to ask for her money back.

Question: Who did Sophie's brother warn Sophie would phone when?

Answer 1: He warned Norman that Sophie would phone on Friday.

Answer 2: He warned that Sophie would phone Mrs Smith on Friday.

Answer 3: He warned Norman on Thursday that Sophie would phone.

All answers are pragmatically true given the story, but syntactic constraints restrict whether they are possible answers or not. Answer 1 corresponds to *who* having scope in the matrix clause and *when* in the embedded clause. There is no syntactic violation; answers 2 and 3 correspond to both *who* and *when* having scope in the embedded clause and the matrix clause, respectively, but violate superiority (*when* being superior to object *who*). Five question types are included (see Table 4.5 below). Each answer chosen by an informant was given a score of 1 and each unchosen answer a score of 0, irrespective of the grammaticality of the answers. An overview of Hawkins and Hattori's (2006) results is given in Table 4.5.

Table 4.5. Mean choice of answers corresponding to the scope of a matrix *wh*-word

	Embedded scope		Matrix scope	
	JSE	NSE	JSE	NSE
Q type				
(a) k=3	0.96	0.97		
SD	0.105	0.100		
(b) k=4	0.78	0.75	0.92	0.91
SD	0.262	0.194	0.119	0.169
(c) k=3	0.75	0.33	0.88	0.85
SD	0.291	0.298	0.165	0.273
(d) k=3	0.58	0.21	0.93	1.00
SD	0.366	0.308	0.140	0.000
(e) k=1	0.58	0.00	0.95	0.91
SD	0.507	0.000	0.229	0.302

Key:

(a) = embedded scope only, no syntactic violation

(b) = matrix and embedded scope, no syntactic violation

(c) = matrix and *embedded scope, violation of superiority

(d) = matrix and *embedded scope, violation of subjacency

(e) = matrix and *embedded scope, violation of superiority and subjacency

Source: Hawkins and Hattori (2006:291)

Here are some of the most striking findings obtained from Hawkins and Hattori's study: *i*) for question types *a* and *b*, where no syntactic violations are involved, JSEs' responses are not significantly different from NSE; *ii*) for question types *c*, *d*, *e*, where violations of superiority or/and subjacency are involved, JSEs' responses are significantly different from NSE; *iii*) JSEs, unlike native speaker controls, do not show significant differences in their acceptance of the grammatical and the ungrammatical readings. These findings support Hawkins and Hattori's prediction that JSEs should not be sensitive to subjacency and superiority violations. This result is in turn interpreted as consistent with the claim that JSEs have failed to represent [*uwh**] feature in C [Q], therefore supporting the IH. How might the JSEs' apparent success (e.g. relating to finding *i*) be accounted for? The authors explain that JSEs have a compensatory representation for the missing [*uwh**] feature, namely a [*uFoc**] feature, which is assumed to be the driving force for scrambling in Japanese. Hawkins and Hattori (2006) assume that in Japanese multiple *wh*-questions involving scrambling, one of the *wh*-phrases must move to value [*uFoc**], but crucially such movement has different properties from *wh*-movement, because scrambling does not give rise to superiority effects. This is illustrated in (4.5):¹⁹

- (4.5) a. John-wa [dare-ga *nani*-o tabeta ka] itta no?
 John-Top who-Nom what-Acc ate C said Q?
 Did John say [who ate what]?
- b. *Nani*-o John-wa [dare-ga <*nani*-o> tabeta ka] itta no?
 What-Acc John-Top who-Nom <what-Acc> ate C say Q?
 What did John say who ate?

Hawkins and Hattori argue that the L1-based [*uFoc**] representation facilitates the JSEs in producing and interpreting English *wh*-interrogatives, but since Focus movement does not involve subjacency and superiority effects, JSEs fail to be sensitive to superiority and subjacency violations. Hawkins terms this study a case of 'features missing despite appearance' (as opposed to 'features present despite appearance'), and cautions that apparent target-like L2 performance may not imply nativelike representation of the L2 grammar.

A problem with this L1-based 'compensatory feature' account is that the current syntactic theory does not provide a solid ground for non-feature-driven movement that Hawkins and Hattori have assumed for Focus movement in Japanese. Their account for Focus movement is based on a series of assumptions and their own speculations, and is therefore, in their own terms, "extremely tentative" (Hawkins and Hattori 2006:297). Pending the refinement of a syntactic theory with Focus movement, their argument that JSEs have a compensatory [*uFoc**] representation for the missing [*uwh**] feature is not fully convincing. A methodological weakness of their study is that they did not include comparative L2 groups (e.g. a L1 Chinese group, whose L1 has a weak [*uwh*] feature, but does not allow scrambling, and/or a L1 French group, whose L1 has the same strong [*uwh*] feature as in English). If these L2 groups turned out to behave in the same way as Japanese speakers, the authors might have to find some other explanations than one

¹⁹ The following symbols are used for annotating examples (4.5*a*, *b*): *Top* stands for topic, *Nom* for nominative case, *Acc* for accusative case, *C* for complementizer, *Q* for question marker.

including compensatory [*uFoc**] feature. Certainly with no such data provided, we will never know how other L2 groups interpret English multiple *wh*-questions. Adding a comparative group, however, would undoubtedly add strength to their study. Lastly, like in Hawkins and Chan's (1997) study, the subjects may not be real endstate L2 speakers of English, so again the claims made about the ultimate attainment are somewhat speculative. In the next section we will review Franceschina (2001), which reports a case study of an endstate L2 speaker, and provides further evidence supporting the FFFH.

4.2.3 Franceschina (2001)

Franceschina (2001) examines spontaneous production data gathered from an adult English speaker of Spanish, named Martin. Martin was first exposed to Spanish at the age of 17. At the time of recording, he had lived in an L2 Spanish immersion situation for a total of 24 years, and was perceived as a highly fluent speaker of Spanish. Franceschina investigates Martin's L2 Spanish number and gender marking on nouns and some syntactic categories, such as pronouns, adjectives and articles and demonstratives.²⁰ She assumes that both Spanish and English have interpretable gender and number features on nouns, but the two languages are parameterized in terms of the uninterpretable counterparts on determiner (D) and adjective (A) heads (see Table 4.6).

Table 4.6. Parameterized uFs in English and Spanish

	D		A	
	[<i>uNUM</i>]	[<i>uGEN</i>]	[<i>uNUM</i>]	[<i>uGEN</i>]
Spanish	+	+	+	+
English	+	-	-	-

Key: + Present in language; - absent in language

Franceschina terms nouns that carry interpretable gender and number features 'triggers', and the syntactic categories that have uninterpretable gender/number features 'targets'. Uninterpretable features on targets must be valued and deleted by the corresponding interpretable features on triggers. Franceschina finds that Martin behaves differently in terms of gender/number marking between triggers and targets: performance on triggers is at ceiling (100%), in contrast to less than target-like performance on target categories (see Table 4.7). A more fine-grained analysis of gender vs. number agreement errors reveals that whether or not Martin made gender/number errors on a certain target depends on the absence or presence of the corresponding feature on that target category. As we can see from Table 4.8, Martin shows a breakdown in gender agreement and number agreement between adjectives and nouns. An additional observation is that Martin uses the masculine form as a default in all gender target categories. Franceschina takes these observations as an indication that the uninterpretable features not instantiated in Martin's L1 (i.e. English) are absent from his L2 syntactic representations. Thus the nature of Martin's problem is L1-based syntactic deficits. Martin's persistent problem with gender/number agreement in his L2

²⁰ Details about Spanish DPs have been presented in Chapter 2 of this dissertation. They will not be repeated here in order to save space.

Spanish thus provides evidence for the FFFH's claim that uninterpretable features not selected into L1s are subject to a critical period.

Table 4.7. Martin's mean scores according to category

Category	Number of obligatory contexts	
Noun	1652	(100.00)
Adjective	577	(92.20)
Article	782	(91.07)
Demonstratives	175	(85.16)
Pronoun	918	(90.49)

Source: Franceschina (2001:236)

Table 4.8. Gender vs. number agreement errors

Category	Gender errors		Number errors	
Adjective	41/53	(77.36)	12/53	(22.64)
Article	65/69	(94.20)	4/69	(5.80)
Demonstratives	48/48	(100.00)	0/48	(0.00)
Pronoun	85/87	(97.70)	2/87	(2.30)
Total	239/257	(93.00)	18/257	(7.00)

Source: Franceschina (2001:237)

A strong point of Franceschina's study is that her subject, Martin, being in an L2 immersion situation for 24 years, really is an endstate L2 speaker. The data obtained are therefore suitable for making claims about ultimate attainment. However, an obvious weakness, which is common to all case studies, is that observations based solely on one subject may not be generalizable. If additional endstate L2 speakers, preferably from different L1 (with or without gender and number features) backgrounds, could be investigated, more convincing evidence would be obtained.

4.3 Full Transfer Full Access (FTFA)

4.3.1 Schwartz and Sprouse (1994, 1996)

We begin with an examination of a case study reported in Schwartz and Sprouse (1994, 1996), where the Full Transfer Full Access (FTFA) hypothesis was first formulated. Schwartz and Sprouse examine longitudinal spontaneous production data (26-month period) gathered from an adult Turkish speaking learner of German, named Cevdet, whose first contact with German was at the age of sixteen. After a year of formal instruction to German, Cevdet's subsequent exposure to German was mostly naturalistic. The recording began shortly after he arrived in Germany. So this study investigates the initial state rather than ultimate attainments of a second language learner.

Schwartz and Sprouse look at Cevdet's acquisition of two properties of German: word order and nominative case assignment. Here we only focus on the acquisition of German word order. German and Turkish VPs are both head-final. The two languages are assumed to be parameterized with respect to the strength of the complementizer (C):

German C is strong, so the finite verb is attracted to C, giving rise to the V2 placement in surface word order. Turkish C, on the other hand, is weak, so the canonical word order of Turkish is SOV. Meanwhile, Turkish also allows scrambling, which enables an XP element (e.g. an adverb, prepositional phrase or direct object) to appear in the sentence initial position. Schwartz and Sprouse counted the following three types of utterances:

- Subject-initial main clauses (SVX)
- Non-subject initial main clauses with pre-verbal subjects (XSV)
- Non-subject initial main clauses with post-verbal subjects (XVS)

Three developmental stages were observed in Cevdet's production data. In stage one, the most common pattern observed is SVX, along with a small proportion (about 22%) of XSV pattern. This was taken as evidence indicating that Cevdet had switched C to the German value, as he consistently moved finite verbs (from clause final position) to C, and subjects to Spec CP. The authors tie the XSV pattern to CP adjunction (i.e. an XP element adjoins to CP), and posit that this pattern in Cevdet's German interlanguage may be due to a transfer of L1 properties related to scrambling. The XSV pattern continues in the second stage. In the meanwhile, a new pattern involving subject-verb inversion (i.e. XVS) emerged during this stage. The XVS should be target-like, but crucially, the inverted subjects are pronominal almost without exceptions. Schwartz and Sprouse point out that the asymmetry between pronominal and nonpronominal subjects in this stage, although not an option in German, can be found in some other natural languages, such as French. The authors then argue that the option adopted by Cevdet in the construction of his L2 German interlanguage grammar is made available by UG. At stage 3, Cevdet had begun to produce target-like German word order (e.g. XVS), where the inverted subjects were not restricted to pronouns. But non-target-like XSV pattern, which was analyzed as involving CP-adjunction, still persisted at stage 3.

Based on the observations about Cevdet's developmental stages in his acquisition of German word order, the authors propose the Full Transfer Full Access Hypothesis (FTFA), which makes the following major claims about L1 transfer and access to UG:

- The entire L1 grammar is transferred into the L2 initial state ("Full transfer" or "absolute L1 influence" in the authors' terms)
- Subsequent UG-based restructuring is possible; the interlanguage grammar is UG-constrained, although not necessarily target-like (Full access).

But if we examine the reported developmental facts in relation to the proposal, two questions arise: first, was the Turkish parameter value regarding the strength of C transferred into the L2 initial state at all? Note that the authors posit that Cevdet had switched C to the German value at the very first stage. Second, when does full access occur? Note that the illicit XSV pattern co-existed with target-like V2 patterns even in the third stage (i.e. the last stage reported in their study). Such a co-existence of V2 and non-V2 patterns is obviously not a UG-derived option. So the question is: can the illicit XSV pattern be eventually abandoned at all?

The authors address the first question by positing an "initial unobserved stage", which they term Stage 0. They hypothesize that during stage 0, C in Cevdet's L2

German interlanguage has the Turkish value, and that the word order he produced would be consistently Turkish-like. Their so doing virtually renders the ‘full transfer’ part of the hypothesis conceptually unfalsifiable: if L1 properties are not observed in the L2 initial state, one can always argue that the L2 data are not collected early enough!

As to the second question, the authors seem not to be in a position to answer it at all, because they obviously did not examine Cevdet’s developmental stages late enough. Nevertheless, they speculate that the XSV pattern is likely to persist into the end stage, as positive evidence indicating the ungrammaticality of such a pattern is presumably not available in German input. Here the authors seem to claim that grammatical restructuring is driven by positive evidence in the L2 input; if the positive evidence is not available, parameter-resetting will not be possible. If grammatical restructuring is forced by positive evidence alone, as the authors claim, what role does UG play in interlanguage grammatical development? Full access to UG entails that learners acquire L2 knowledge without positive evidence, as pointed out by Lardiere (2006a:70): “If there were positive evidence available in the input, there might be no reason to invoke UG at all, anyway. One need invoke full access to UG only in the absence of positive evidence.”

Despite the above mentioned weakness of Schwartz and Sprouse’s study, the core ideas of the FTFA model they proposed are appealing, and have been adapted to make more testable hypotheses. For example, White (2003:61) claims that two types of evidence serve to support FTFA (i) evidence of L1 properties in the interlanguage grammar; (ii) evidence of restructuring away from the L1 grammar.

It is clear from the above discussion that the FTFA model originally was not intended to be a theory about ultimate attainment. Nevertheless, it has been extended to make predictions on the final outcomes of L2 grammars, and it has been reinterpreted in minimalist terms in response to the advances of general linguistic theory, as in the study by White *et al.* (2004) to which we will turn in the next section.

4.3.2 White *et al.* (2004)

White and her colleagues defend the FTFA against the FFFH, claiming that the full set of abstract features provided by UG remains available to L2 learners, regardless of age of acquisition. Contra the FFFH, according to which L2 grammars cannot represent parameterized uninterpretable features, White *et al.* argue in accordance with FTFA that “nativelike mental representations are in principle acquirable” (White *et al.* 2004:106). They further offer an account for variable gender agreement in nonnative (and native) speakers, following the same line of reasoning as Lardiere (1998a, 1998b), who argues for ‘dissociation between syntax and morphology’, and as Prévost and White (2000), who propose a *Missing Surface Inflection Hypothesis* (MSIH). This line of argument is altogether referred to as a *morphological approach* in Franceschina (2001).

The FTFA and the FFFH were evaluated in relation to the acquisition of gender and number agreement in L2 Spanish by speakers of English (which does not have grammatical gender), and speakers of French (which has grammatical gender). Subjects are at three proficiency levels: low, intermediate, and advanced. However, Spanish was an L2 only for 14 (out of 68) of the English-speaking learners; for others (n=54) it is an L3, the L2 being French (among them 36 had first contact with French at the age of 9 or under; 18 at the age of 10 or above). A DP analysis similar to that in Franceschina (2001, see Section 4.2.3) is assumed. That is, adjective (A) and determiner (D) heads in

Spanish carry uninterpretable gender and number features to be ‘checked’ by the corresponding interpretable features on the noun (N). French has exactly the same parameter settings as Spanish, while English and Spanish are parameterized with respect to uninterpretable gender on A and D. Two tasks are used to elicit DPs containing adjectives. One is a ‘Guess who’ game, the other is a picture description task. In accordance with the FTFA, White *et al.* made the following predictions:

1. a. At lower proficiency levels, learners whose L1 is English will perform more accurately on number agreement than on gender agreement.
- b. At lower proficiency levels, learners whose L1 is English will have greater difficulty with gender agreement than learners whose L1 is French
2. a. Advanced learners should perform as accurately on gender as on number
- b. Advanced learners should perform accurately on gender agreement, regardless of the status of gender features in the L1

They also include a comprehension task, exploiting a property in Spanish known as noun-drop (N-drop), where the head noun in a Spanish DP can be null provided that its content is recoverable from the context. In the comprehension task, subjects are asked to read a story consisting of 48 sentences, each including a DP with a null nominal. Each sentence appears with three pictures. The selection of the appropriate picture depends on the gender and number cues available on adjectives and/or determiners in the DP. If uninterpretable gender features are not available, White *et al.* predict that the English group, regardless of proficiency levels, cannot consistently interpret the null nominals identified by gender. The main findings from White *et al.*’s study are listed below:

- Number is unproblematic for either group at any proficiency levels (partially consistent with prediction 1a);
- At lower proficiency levels, gender is problematic for both the English and the French groups (contra prediction 1b);
- At advanced proficiency levels, both groups are as accurate on gender as on number (in accordance with prediction 2a), and the English group is as accurate on gender as the French group (in accordance with prediction 2b);
- As for gender-marking, both learner groups show greater accuracy on feminine items (in other words, masculine surfaces as a default);
- For the English-speaking learners, there is no significant difference between those who were earlier exposed to French and those who had no prior exposure to French;
- Results from the comprehension task are consistent with those of the production tasks.

The observed similar acquisition patterns exhibited by the English-speaking and the French-speaking learners lead the authors to conclude that the absence of gender in the L1 is not the only factor affecting L2 gender acquisition, contra the FFFH. The eventual ultimate convergence of L2 grammars with native grammars is interpreted as evidence that the uninterpretable gender features on determiners and adjectives have been successfully acquired, supporting the FTFA. A potential weakness of this study is that the many subjects in the English group had exposure to a language with gender before

puberty. So it is difficult to rule out the possibility that the uninterpretable gender features had been acquired before a critical period. Though performance between the early exposure group and the no exposure group has been compared and no significant difference is found, no information is provided about proficiency levels of the subjects in each group.

At the end of the paper, White *et al.* offer an account for variable gender agreement, which is taken to be a problem common to both nonnative speakers and native speakers (e.g. the variable gender assignment in 19th century Quebec French being 7%, citing Klapka (2002)). The variability in general is accounted for with a reference to *Distributed Morphology (DM)*, (See for example Halle and Marantz 1993), according to which feature checking in the syntax involves movement of abstract bundles of features, not actual lexical items, and that lexical items undergo late insertion into the structure. Crucially under this view, features in the syntax are fully specified, whereas those on lexical items can be underspecified, so far as no feature clash result. In the case at hand, masculine is the default, which is assumed to be deprived of any feature values; feminine is the only value to be entered in the lexical entry of Spanish nouns. Under these assumptions, the insertion of gender features proceeds in the following way:

If a noun is marked [+fem], the determiner and adjective positions will become feminine by agreement (or feature checking), so feminine forms can be inserted. However, masculine forms of determiners and adjectives (not specified, hence defaults) are not precluded because they do not result in a feature clash: feminine in the syntax, no specification on the lexical items. On the other hand, if the head of the DP is marked [+masc], only default forms unspecified for gender can be inserted, because insertion of [+fem] items would result in a clash of features...In the normal course of events, the most specified forms “win” as far as lexical insertion is concerned, which is why failure of gender agreement is unusual in native speaker grammars and why the learners show correct gender agreement in the majority of cases (White *et al.* 2004:129).

As such, White *et al.* are arguing in the same line of reasoning as Lardiere (1998a, 1998b) and Prévost and White (2000), who propose that there is a dissociation between syntax and morphology in second language acquisition, and that errors committed by L2 speakers due to mapping problems between surface forms and abstract features, rather than deficits in syntactic representations. These authors argue for the modularity of grammatical subsystems (e.g., knowledge of phonology) and the existence of the interface “translation” procedures for mapping one area to another. They contend that it is the interface between the morphological module and the syntactic module (rather than the syntax) that is affected by a critical period (see for example Lardiere 2006a:235-236). Lardiere (2006a), for example, claims explicitly that adults differ from children “in the process of working out the correspondence between abstract features and phonological forms” (p. 93). Her claim is based on her observations of the ultimate attainment of an endstate adult Chinese speaker of English, named Patty. Through examining Patty’s longitudinal production data, Lardiere (2006a) finds that Patty’s suppliance of tense and verbal agreement morphology is at a very low rate across the board, in contrast to her perfect suppliance of nominative case marking on subjects in finite clauses. This finding is interpreted as evidence for successful acquisition of finiteness features associated with T(ense) head, because it is assumed that a nominative case feature of a subject DP is checked within the domain of T[+finite]. What is problematic for Patty is the mapping or translation procedures from abstract syntactic features to the morpho-phonological spell-out of inflection. In a similar vein, Prévost

and White (2000) propose that a mapping problem between surface forms and abstract features is responsible for the missing surface inflection in L2 performance. This proposal is referred to as the “missing surface inflection hypothesis.”

This line of reasoning (which Franceschina calls a *morphological approach* to explaining NNS/NS divergence) is under scrutiny in Franceschina (2001), who discusses both theoretical and empirical problems for this approach. At the theoretical level, the adoption of Distributed Morphology (DM) as a model for the syntax-phonology interface is called into question. First, DM’s approach to lexical insertion is at odds with that of the MP: post-syntactic (or late insertion) in DM, pre-syntactic in the MP. Second, DM posits that there exists a level of morphological structure which constitutes the interface between the syntax and the phonology, whereas in the MP, only the sensory-motor interface (SM) and the conceptual-intentional interface (C-I) are considered to be “optimal properties” (cf. Chomsky, forthcoming). It is not clear how the two models are made compatible in the morphological approach. An empirical problem, pointed out by Franceschina, is that the morphological approach cannot explain why the morphological reflexes of Case have been acquired, while the reflexes of tense and agreement have not. As Franceschina comments, “it is at least dubious that the syntax-morphology mapping mechanism should falter when dealing with past form but not with case-marked forms.” She further points out that the morphological approach fails to explain why her subject, Martin (see Section 4.2.3 above), has problems with genders on target categories (adjectives and determiners), but not on triggers (i.e. nouns), as the morphological forms are the same across all categories marked for gender in Spanish.

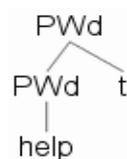
Variability in production of morphology in L2 performance is indeed intriguing. It stimulates L2 researchers to look for new approaches in order to arrive at more articulated accounts. One of the recent and most prominent proposals is the *Prosodic Transfer Hypothesis* (Goad *et al.* 2003; Goad and White 2004; Goad and White 2006). We will look at that proposal in the section to come.

4.3.3 Goad *et al.* (2003), Goad and White (2006)

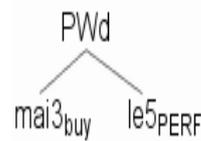
Goad *et al.* (2003) propose the *Prosodic Transfer Hypothesis* (PTH), according to which L1 prosodic constraints restrict the types of representations that can be built in the L2, hence limiting IL production of inflectional morphology and functional words. Particularly, the PTH argues that L2 representational deficits are situated in the phonological component rather than in syntax. Accordingly, it claims that L2 learners’ underlying knowledge of the L2 morphosyntax is nativelike, but their realization of overt morphology is constrained by L1 prosodic representations; the L1-based prosodic representation may persist in the end state of L2 acquisition. It is immediately clear that the PTH is compatible with the FTFA model. In accordance with the PTH, syntactic features are fully available to L2 learners (*full access*); L1 prosodic structures (and presumably other L1 properties) are transferred into the L2 and constrain L2 grammatical representations (*full transfer*). Though apparently arguing against the Representational Deficit Hypothesis (RDH), the PTH has one point in common with the RDH, that is, both hypotheses claim that L2 competence is constrained by L1 representations. But they differ as to the source of L1 effects: phonological in the case of the PTH; syntactic in the case of the RDH (see Goad and White 2006:246; Lardiere 2006a:210).

Goad *et al.* (2003) look at the acquisition of English tense and agreement morphology by Chinese-speaking learners. They assume that the inflectional morphology in the two languages is prosodified differently. In English, regular inflection is adjoined to a prosodic word (PWd), as shown in (4.6a) for *helped*. As such, functional and lexical materials are represented differently from each other in both components of the grammar; in Chinese, by contrast, inflection is incorporated into the PWd as an “internal clitic”. This is illustrated by *mai* ‘buy’ with a perfective (PERF) aspect *le* in (4.6b). Consequently, prosodic structure does not reflect syntactic organization: functional and lexical materials are prosodically formally equivalent. However, the prosodic structure of irregular inflection in English (including pseudo-inflection in (4.6ci) and ablaut in (4.6cii)) resembles that of Chinese, that is, it is organized PWd-internally.

(4.6) a. English regular inflection

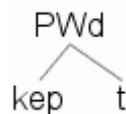


b. Mandarin inflection



c. English irregular inflection

(i). Pseudo-inflection



(ii) Ablaut



(Illustrations from Goad and White 2006: 247-245)

Goad *et al.* (2003) examine the production of English tense and agreement morphology by 12 Mandarin-speaking adult learners of English, whose proficiency level ranges from high intermediate to low advanced. They predict two patterns of behavior in accordance with the PTH:

1. Across-the-board (ATB) deletion: Learners have come to realize that English does not permit a stem-internal analysis of inflectional morphology and are sensitive to the need for a unified analysis of this morphology. However, their grammars do not permit adjunction in outputs. The result is ATB deletion of inflection.
2. Variable deletion of inflection: Inflection morphology surfaces for stimuli where it can be incorporated into the PWd (like Mandarin aspect *-le*). The result is variable deletion of inflection but the variability is predictable.

Just as they predicted, subjects turned out to fall neatly into two distinct groups: an ATB deletion group (suppliance of *-s* at the rate of 10%) and a variable deletion group (suppliance of *-s* at the rate of 49%). They further posit three conditions under which a PWd-internal analysis of inflection is possible: (a) agreement as onset, for example, ‘*builds on*’; (b) agreement as coda, for example, ‘*races*’; (c) agreement as foot-internal onset of an empty-headed syllable (OEHS), leading to a structure parallel to that attested for Mandarin aspect, for example, ‘*fills*’. They found that the variable deletion group

supplied *-s* at a much higher rate in the three situations than when there is no option for agreement to be inside the PWd, as shown in Table 4.9.

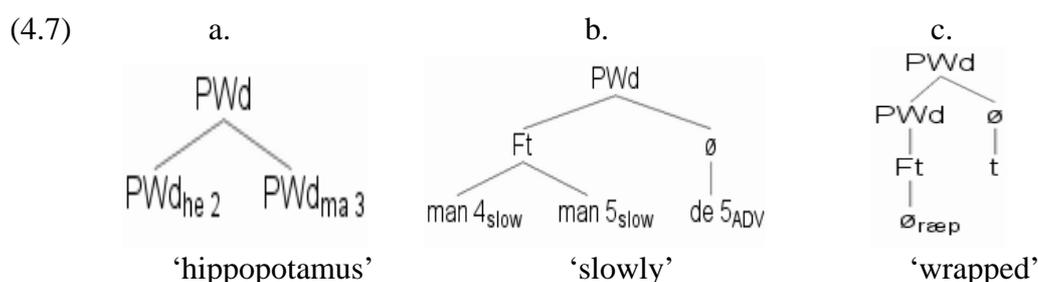
Table 4.9. Agreement (Agr) in production by stem shape (% target-like)

	ATB deletion group	Variable deletion group
Agr as onset	7	75
Agr as coda	0	27
Agr as foot-internal OEHS	7	68
No option for Agr inside PWd	0	9

Source: Goad *et al.* (2003:259)

Moreover, they observe that subjects are perfect in nominative case assignment (100%), and accurate in supplying copula (97%) and auxiliaries (87%). This observation is interpreted as evidence indicating that subjects represent tense and agreement features in their interlanguage grammars. Omission of inflection then is attributed to L1-based prosodic representations. The authors further posit that L1 prosodic representations may be subject to fossilization. As a consequence, target-like prosodic representations are predicted to be unattainable.

Goad and White (2006) provides further evidence supporting the PTH. However, they modified its claim about ultimate attainment, contending instead that “target-like prosodic representations are ultimately attainable for at least some functional material which is absent from the L1”(p. 264). On the assumption that the adjunction prosodic structure is available in English, but absent in Chinese, Goad and White predict that supplience of regular tense and participle morphology will be depressed. However, they hold that prosodic representations for lexical material in the L1 can be adapted to represent functional morphology in the interlanguage, and that target-like prosodic representations can be built from existing L1 structures. According to them, the necessary adjunction structure can be built by combining lexical compounds, as in (4.7a), and three-syllable PWds, as in (4.7b), available in Mandarin. The result will be the target-like adjunction structure for English regular inflections, as in (4.7c).²¹



This prediction is tested through systematically investigating the acquisition of English tense and participle morphology by 10 intermediate Chinese-speaking learners. Subjects are tested using a sentence completion task, where they are first asked to choose an answer from pairs of sentences with contrasting morphology (*present vs. past*; *present vs. perfective*), and then speak their chosen ending aloud. Stimuli contain regular and irregular verbs, and according to stem shape, regular stimuli are further

²¹ The symbol ‘Ft’ stands for foot, ‘∅’ for syllable.

divided into long stem (e.g. *helped*) and short stem (e.g. *wrapped*); irregular stimuli into pseudo-inflected (e.g. *kept*) and ablaut (e.g. *won*). Results show that subjects are accurate with the sentence completion task (87% accuracy or above), and there are no differences between past and perfective forms in production. This finding is interpreted as indicative of nativelike representations at the syntactic level. A fine-grained analysis of the articulatory properties of the stem consonant reveals an asymmetry in the distribution of fortis release: it occurs in short(ened) regulars, but not in pseudo-inflected forms. This differential treatment of regular and pseudo-inflected forms is claimed to suggest that the two types of inflections are represented differently by L2 speakers: adjunction for regular forms, whereas PWd-internal structure for pseudo-inflected forms. In other words, the adjunction structure (which is claimed to be absent in Chinese) has been successfully acquired by Chinese-speaking learners. The authors then conclude that interlanguage grammars are *initially* constrained by L1 prosodic representations, but target-like representations are acquirable “through combining licensing relations available in the L1” (p. 265).

If we compare the results from the two studies (i.e. Goad *et al.* 2003; Goad and White 2006), the question arises of why there should be such a big discrepancy in the suppli-ance of English inflections between the two groups of Chinese-speaking subjects, given that their English proficiency is at equivalent levels?²² Goad and White (2006) attribute this to task differences, contending that the sentence completion task used in Goad and White’s (2006) study may make subjects “more inclined to realize the inflection” (p. 255). However, if the learners have acquired target-like representations (i.e. the adjunction structure) as the authors have assumed, it is at least dubious why their performance should be so much influenced by task differences. It should be recalled that in the former study (i.e. Goad *et al.* 2003), half of the subjects delete tense and agreement morphology across the board; the other half supplied them at chance level. It is not clear why particularly these learners cannot build the appropriate representations from the existing L1 structures.

Moreover, the (picture description and sentence completion) tasks used in the two studies to some extent allow learners to draw on metalinguistic knowledge. Hence it is problematic to interpret the subjects’ nativelike performance as indicative of nativelike grammatical knowledge (about tense/agreement). As the authors acknowledge, they simply reflect learners’ knowledge at a *metalinguistic* level. Therefore, results from the two tasks do not provide compelling evidence against the RDH. A more general problem with the two studies is that the data are limited with respect to L1/L2 combinations and the kinds of learners being tested. As we have seen, both studies focus on the acquisition of English as the second language, where inflections are realized as consonants, and both test Chinese-speaking learners. Clearly, the PTH is in need of further testing. More systematic investigation involving a variety of target languages and comparing inflections of different phonetic types (i.e. consonants vs. vowels, as in the current study where inflections involve three different phonetic types: *-t* (consonant), *-e* (vowel), *et/en* (vowel + consonant)) is required in order to make it more compelling.

²² Actually some subjects in the former study are at the even higher level than the subjects in the latter one.

4.4 Shallow Structure Hypothesis (SSH)

The shallow structure hypothesis (SSH) was proposed by Clahsen and Felser (2006a) in a keynote article that appears in a special issue of *Applied Psycholinguistics*. This issue also includes a number of commentaries on the target article, and the authors' response (i.e. Clahsen and Felser 2006b) to these comments and questions. Clahsen and Felser (2006a) present a comprehensive overview of the empirically-based psycholinguistic studies, from which the authors generalize their findings about child L1 processing and adult L2 processing. They argue that children behave basically like adults (*Continuity Hypothesis*), but that adult L2 learners behave in a fashion that is fundamentally different from native speakers. To account for the claimed L1/L2 differences, they propose the Shallow Structure Hypothesis (SSH), according to which the representations adult L2 learners compute during processing contain less syntactic detail (which are defined as representations without hierarchical syntactic organization) than those of child and adult native speakers.

It is important to note that the theory differentiates between relatively simple morphological rules and the computation of complex syntactic representations. For morphological processing, L2 learners are claimed to be able to employ the same mechanism as native speakers; for sentence processing, however, L2 learners are claimed to rely more on lexical-semantic ("shallow") cues rather than syntactic information. Evidence supporting this argument comes respectively from L2 learners' processing of inflected words, and their processing of ambiguous sentences and syntactic dependencies. In the following, we look at one representative study in each area, namely Hahne *et al.* (2003) and Marinis *et al.* (2005).

Hahne *et al.* (2003) use event-related brain potentials (ERPs) to investigate how adult Russian-speaking learners of German process German participle formation and noun plurals on line. In both cases, participants were presented with sentences containing two types of violations, overregularizations (formed by adding a regular suffix *-t/-s* to a verb/noun that requires an irregular one) and incorrect regular forms (in which a verb/noun that takes the regular default suffix *-t/-s* appeared with a different incorrect ending). Brain responses of the L2 learners are compared with native German speakers in the previous studies (Penke *et al.* 1997; Weyerts *et al.* 1997; Lück *et al.* 2001). It was found that advanced L2 learners do not differ fundamentally from native speakers in how they process morphologically complex words, as can be observed from Table 4.10.

Table 4.10. ERP effects to morphological violations in L2 and native speakers

	L2 learners	Native speakers
INCORRECT REGULARS		
Participles	N400	No effect
Plurals	N400	N400
OVERREGULARIZATIONS		
Participles	LAN,P600	LAN
Plurals	P600	LAN,P600

Source: Clahsen and Felser (2006:15)

Before we proceed with the authors' interpretation of the results, a brief explanation of the technical terms is in order. LAN stands for left anterior negativity. It reflects the automatic first-pass parsing process (such as detecting morphosyntactic mismatch); centroparietal N400 is associated with *lexical-semantic* processing. The P600 stands for centroparietal positivity. It is a component which indexes the process of *syntactic* reanalysis and repair (cf. Clahsen and Felser 2006a). Results from Table 4.10 show that L2 learners, like native speakers, respond differently to violations of regular and irregular rules of inflection. For irregularizations, they showed N400 effects, which are indicative of lexical storage of irregulars. For overregularizations, they showed P600 effects, which are indicative of morphological decomposition (i.e. the unmarked stem + the *-s* plural/*-t* participle rule). The only difference between L2 learners and native speakers, however, is in terms of LAN for noun plurals, which is found for native speakers but not for L2 learners. The authors attribute this difference to the L2 learners' relatively insufficient mastery of noun plurals. In an elicited production task after the ERP experiments, they found that the L2 learners are less accurate with plural formation than participle formation (86% vs. 95%) for critical items used in the ERP experiments. The authors therefore conclude that L2 learners, be they proficient enough, do not fundamentally differ from native speakers in morphological processing.

Turning to L2 processing in the syntactic domain, Marinis *et al.* (2005) carry out a self-paced reading study investigating how L2 learners of English process sentences involving long-distance *wh*-dependencies of the kind shown in (4.8 a-b).

- (4.8) a. The nurse *who* the doctor argued ___ that the rude patient had angered ___ is refusing to work late. (+ intermediate gap)
 b. The nurse *who* the doctor's argument about the rude patient had angered ___ is refusing to work late. (- intermediate gap) (examples from Clahsen and Felser 2005:36)

Following Chomsky's copy theory of *wh*-movement (1995), the authors assume that an intermediate gap is present in (4.8a) at the intervening clause boundary, breaking the long dependency up into two shorter ones, whereas no such intermediate gap is present in (4.8b). Experimental sentences were presented visually in a segment-by-segment fashion, and were followed by a comprehension question (e.g. *Who angered the nurse?*). The authors predict that participants who postulate intermediate gaps should find it easier to integrate the filler with its subcategorizer in sentences of type (4.8a) than in sentences of type (4.8b), which should be reflected in shorter reading times on the segment containing the subcategorizing verb (=segment 5) in condition (4.8a) than in condition(4.8b). Participants include four groups of L2 learners with similar levels of proficiency in L2 English, two from *wh*-movement language backgrounds (German and Greek), two from *wh-in-situ* backgrounds (Chinese and Japanese), and a group of native speaker controls. Results are presented in Table 4.11.

As we can see, only the native speakers show a significant interaction between Extraction and Phrase type on segment 5, indicating that the presence of an intermediate gap facilitates filler integration only for the native group, but not for L2 learner groups from either (*wh*-movement vs. *wh-in-situ*) L1 backgrounds. The authors interpret this finding as evidence indicating that native speakers consult a mental representation of the filler *who* (referring to the nurse) when they encounter the complimentizer *that* which

signals the beginning of a new subordinate clause, in accordance with the subadjacency principle, whereas L2 learners, irrespective of the presence of *wh*-movement in their L1s, do not observe subadjacency during processing, but try to establish a direct link between the filler and its subcategorizer in both extraction conditions.

Table 4.11. Response accuracy and effects to *wh*-dependencies in L2 learners and native speakers

Participants	Response Accuracy	Effects on Segment 5	
		Extraction Effect	Extraction×phrase type interaction
Native speakers	79.5%	yes	yes
L1 Greek	79.75%	yes	no
L1 German	84.75%	yes	no
L1 Chinese	79%	yes	no
L1 Japanese	74.5%	yes	no

Source: Clahsen and Felser (2006:44)

On the assumption that L2 learners are able to associate a displaced element with its subcategorizer or thematic role assigner, the authors offers a shallow structure account of how sentence (4.8a) might be segmented and analyzed semantically (cf. 4.9a-c, which is 5a-c in Clahsen and Felser 2006a):

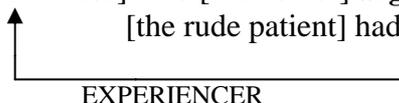
(4.9) a. [The nurse] who [the doctor] **argued** [that...



b. [The nurse] who [the doctor] **argued** [that...
[the rude patient] had **angered**...



c. [The nurse] who [the doctor] **argued** [that...
[the rude patient] had **angered** is refusing to work late.



The presentation thus assembled can lead to nativelike comprehension, but does not contain any intermediate gaps, explaining why intermediate gap effects are absent in L2 processing. Given that L2 processing is generally shallow in the sense described above, the lack of L1 transfer effects is expected. The authors further argue that shallow structure processing is an option available to the human language comprehension system in principle. But contrary to native speakers, “adult L2 learners are largely restricted to this option in L2 processing, computing representations for language comprehension that lack syntactic detail, and attempting more direct form-function mapping instead” (Clahsen and Felser 2006a:62).

Note that in the target article, Clahsen and Felser do not mention where morphosyntactic agreement should fall in the morphology-syntax dichotomy they have stipulated. In the response article, the authors claim explicitly that morphosyntactic

agreement, like word segmentation, belongs to “local” domain phenomena, and that nativelike processing is predicted to be possible. But the rationale for putting morphosyntactic agreement into the domain of morphology is unclear. Although I agree with the authors in their argument that inflected words “involve the simple concatenation of two adjacent elements (stems and affixes) and thus have a much shallower internal structure than sentences” (p. 61), I do not think that the internal structure for morphosyntactic agreement is as shallow as inflected words, as the authors seem to claim. Under current minimalist assumptions, both morphosyntactic agreement and “non-local” (e.g. subject-verb) agreement involve the checking of formal features between two constituents that are in a hierarchical syntactic structure (i.e. in a c-command configuration). It is not obvious why formal feature checking is possible for morphosyntactic agreement, but impossible for non-local agreement. Moreover, if ‘local’ means two elements being adjacent to each other, morphosyntactic agreement between D and N is not ‘local’ at all due to the existence of several intervening heads in accordance with Julien’s (2003, 2005) DP model.

If L2 processing is “fundamentally” different from L1 processing, as the authors claim, what could be the possible source(s) of L1/L2 processing divergence? Clahsen and Felser (2006b) suggest two possibilities:

One possibility is that the same parsing mechanisms that are used in L1 processing are also available in L2 processing, but that their application is restricted due to the knowledge source that feeds the structural parser, the L2 grammar, being incomplete, divergent, or of a form that makes it unsuitable for parsing. The second possibility is that although the L2 grammar is sufficiently detailed and suitable for parsing, full parsing fails due to the unavailability or deficiency of the required parsing mechanism (Clahsen and Felser 2006b:117).

The authors believe that the first possibility is more “realistic”, on the assumptions that the parsing mechanisms are universal, whereas L2 grammars are fundamentally different from L1 grammars. Implicated in this argument is the view that the observed L1/L2 processing differences do not merely reflect differences at the level of processing, but also at the level of grammatical representation. That is to say, the source of non-nativelike L2 processing is inadequacies of L2 grammar. Based on this view, the authors predict that L2 learners can develop nativelike parsing abilities when acquiring a nativelike grammar. Note, however, that the grammar-parser relationship assumed by the authors is still controversial and is in need of further testing. In order to test whether nativelike parsing abilities are indeed dependent on acquiring a nativelike grammar, it will be useful to test those L2 speakers who have acquired nativelike grammatical knowledge. Existing empirical psycholinguistic studies have mainly targeted on advanced L2 learners, but rarely on endstate L2 learners who have attained nativelike grammatical knowledge. It remains to be known whether these groups of L2 learners will process the target language in the same fashion as native speakers.

In short, SSH is in need of further elaboration and empirical testing. As the authors note, a comparative approach, involving different groups of language learners, using a variety of experimental method, will turn out to be useful in this regard.

4.5 A summary

In this chapter I have presented an overview of the studies that respectively provide evidence supporting two models on L2 grammatical competence, namely, the Failed Functional Features Hypothesis (FFFH) and the Full Transfer Full Access Hypothesis (FTFA). Figure 4.1 summarizes the two models' respective accounts for the eventual outcomes (convergence vs. divergence) of L2 acquisition. I also have reviewed one model of L2 processing, namely, the Shallow Structure Hypothesis (SSH). The core ideas of the SSH is illustrated in Figure 4.2 (based on Clahsen and Felser 2006b:119).

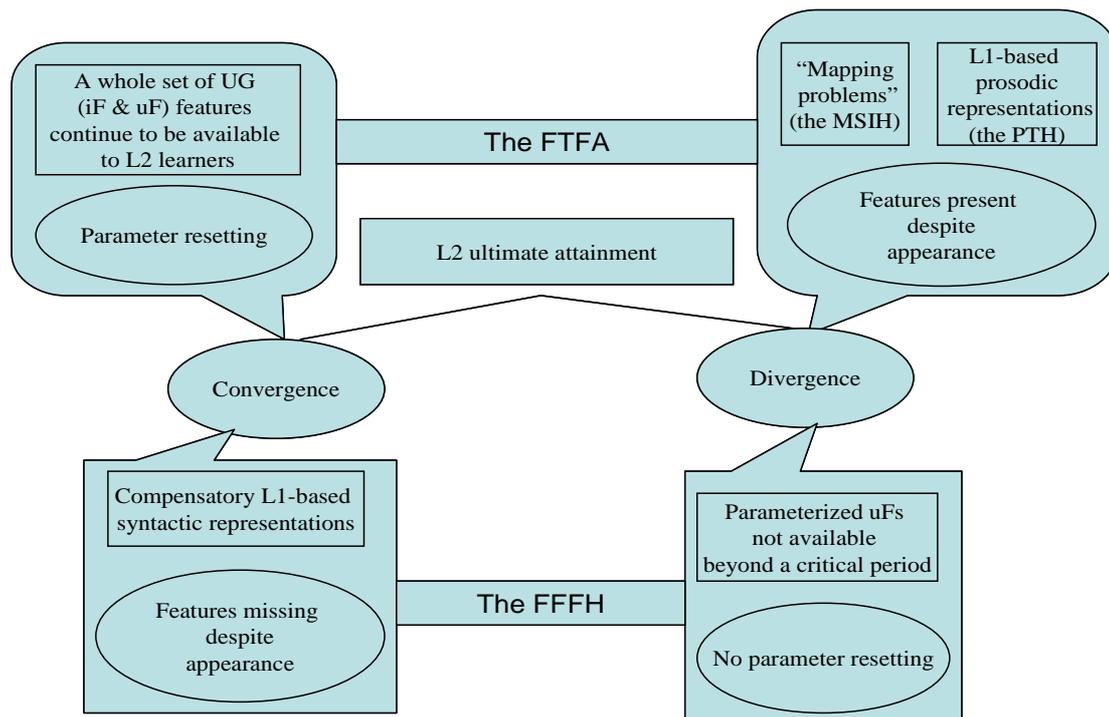


Figure 4.1. The FFFH's and the FTFA's accounts for L2 ultimate attainment

As we can see from Figure 4.1, the FFFH and the FTFA have different approaches when it comes to accounting for convergent/divergent outcomes of adult L2 acquisition. The FFFH, assuming that the parameterized uninterpretable features are affected by a critical period, claims that the interlanguage representation is permanently defective due to the absence of the novel uninterpretable features required to represent the L2. Apparent target-like L2 performance is attributed to the existence of compensatory L1-based syntactic representations. As such, the FFFH falls within the no-parameter-resetting and features-missing-despite-appearance camps. The FTFA, on the other hand, maintains that the whole set of UG formal features continue to be available to adult L2 learners. Thus it holds that interlanguage grammars can converge on target L2 grammars in the syntactic domain. To account for the observed divergent outcomes, the FTFA works out two proposals, namely, the MSIH and the PTH. According to the MSIH, the missing morphology in L2 production is due to problems as to mapping between surface forms and abstract features, rather than deficits in syntactic representations. In a similar vein, the PTH argues that interlanguage representations are

‘perfect’ in the syntactic domain, but are deficit in the phonological domain in that they are constrained by L1 prosodic representations. So contrary to the FFFH, the FTFA falls within the parameter-resetting and features-present-despite-appearance camps.

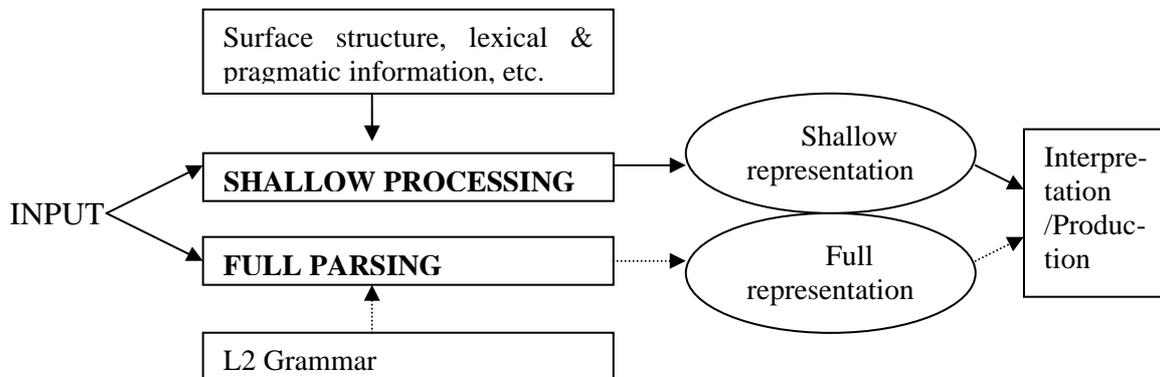


Figure 4.2. L2 processing model in accordance with the SSH

As illustrated in Figure 4.2, the SSH claims that of the two different routes made available by the human language processing system, the shallow processing route, which is guided by semantic/pragmatic information and strong associative meaning or form patterns, is predominant in L2 processing; the full parsing route is available to L2 learners in principle, but it is underused in L2 processing due to inadequacies of the L2 grammar.

It seems that neither the FFFH nor the FTFA share the viewpoint underlying the SSH that L2 grammars are “fundamentally” different from L1 grammars. The FFFH, though arguing that L2 grammars cannot represent those uninterpretable features not selected into the L1, would predict that L2 grammars may converge on L1 grammars should the target language contain no novel uninterpretable features. Convergence (at least at the levels of syntactic presentations) is also predicted for L2 ultimate attainment in accordance with the FTFA. If this is so, following the line of reasoning of the SSH, nativelike L2 processing is predicted to be possible in the areas where L2 speakers have developed nativelike grammatical knowledge. If, on the other hand, L2 processing still turns out to be different from L1 processing, the basic assumptions of the SSH that the parsing mechanisms are universal, whereas L2 grammars are fundamentally different from L1 grammars, will be called into question.

5 *L2 production*

5.1 Introduction

In this Chapter, L2 speakers' grammatical knowledge of Norwegian with respect to DP internal agreement will be investigated using elicited spoken production data. The previous chapter has shown that the functional heads of Norwegian D and α carry unvalued number, gender, and definiteness features. The primary aim of this chapter is to assess the status of the formal features in endstate L2 learners' interlanguage grammar. Three groups of L2 learners; the functional heads D and α of their L1s differ parametrically from the target language with respect to one, two or all of the uninterpretable features. The results from the three groups of learners can therefore provide evidence as to the availability of parameterized uninterpretable features to post-critical period L2 speakers, and can ultimately reveal whether the endstate L2 speakers can develop nativelike grammatical representations. As we have seen, this issue has been the subject of considerable debate in L2 research, especially between the failed functional features hypothesis and full transfer full access model. Where it is necessary, data from the experiments will be analyzed both at the group level and at the individual level, because we are interested in the grammars of individual speakers as well as of groups.

This chapter is organized as follows. Section 5.2 presents detailed background information, inclusion criteria, scores of Norwegian proficiency test of the L2 participants. This is followed by some brief information about the L1 participants who serve as native controls. Two experimental studies are reported in the subsequent two sections. The first experiment, presented in Section 5.3, is an off-line gender assignment task. This task targets gender. Through this task, we wish to find out whether the L2 speakers can associate correct gender with Norwegian nouns. The second experimental study, reported in Section 5.4, is an on-line DP production task, through which the L2 speakers' knowledge about various aspects of Norwegian DP internal (gender/number/definiteness) agreement will be tested. Predictions based respectively on the FTFA and the FFFH are made. The results from the second experiment are presented in Section 5.5, where the informants' suppliance of gender for the same nouns in the two experiments will also be carefully examined in order to find out whether they are consistent in gender-marking. Finally, in Section 5.6, the findings from the experiments are summarized, the predictions of FFFH and FTFA are revisited and

evaluated based on the observations from the experiments, and the L2 data are discussed in comparison with data from first language acquisition studies in literature.

5.2 The participants

5.2.1 The L2 participants

Seventeen L2 speakers of Norwegian (8 males, 9 females) were initially recruited. They came from three typologically different language backgrounds: Italian/Spanish, English, and Chinese. Since Italian and Spanish are alike with respect to DP internal agreement (as has been previously discussed), informants from the two language backgrounds were taken as one group, and referred to as the Romance group. The letter R will be used to identify them. Similarly, English and Chinese informants will be identified by letter E and C, respectively. Six English informants, six Chinese informants, and five Italian/Spanish (2 Italian; 3 Spanish) informants were selected according to the following criteria:

- i) They had to have had residence in Norway for at least eight years;
- ii) They acquired Norwegian as adults (>13 years old), and did not learn any other Scandinavian languages before age 13;
- iii) Their self-rated oral comprehension and oral production of Norwegian were both over 8, on a 10 point scale;
- iv) For English-speaking informants, they did not acquire any gender-marking languages as a child (ages from 0 to 13);
- v) For Chinese speaking informants, they did not acquire any gender or number marking languages as a child (ages from 0 to 13).

All the informants were interviewed by the author regarding their biographical information, L1 and L2 language learning experiences. They were also asked to self-rate their Norwegian oral comprehension and oral production on a ten-point scale (1=very poor; 10=excellent). Details of the informants' information and their self-rated proficiency levels are given in Table 5.1 below. The two Italian-speaking informants (R1 and R2) were from Italy. Two of the Spanish informants (R3 and R4) were from Spain; the other one (R5) came from Venezuela. All the English-speaking informants were from the United States, except for E3, who came from England. Among the Chinese informants, five were from Mainland China, and one (i.e. C6) was born in Mainland China but grew up in Taiwan; all acquired Mandarin Chinese as their mother language. All the L2 informants had received their bachelor's degree or above. So it was assumed that they had comparable intellectual capacities. All had received at least one-year formal classroom instruction in the Norwegian language in the initial years after they migrated to Norway. All of them had become fully integrated in local Norwegian communities through marriage, work or education. Out of the 17 informants, 12 of them (6 English, 3 Italian/Spanish, and 3 Chinese) had a Norwegian Spouse; 9 of them had their university degrees from a Norwegian university; and all of them had worked or studied in a fundamentally Norwegian-speaking environment. They all had great motivation to acquire Norwegian, and appeared to be very successful, as can be seen

from their reported near perfect Norwegian oral comprehension and production (≥ 8 on a 10 point scale for all informants). All of them reported that they were perceived as fluent speakers of Norwegian by the local people, but nevertheless their accents were distinguishable from those of native speakers, especially in the way they pronounced the Norwegian /l/ and /r/.²³

Table 5.1. L2 informants' background information and self-rated language proficiency

ID	M/F	AGE	LOR	ASSN	NOC	NOP
R1	M	44	11	29	9	9
R2	F	40	12	24	9.5	10
R3	M	56	20	32	9	9
R4	F	53	25	28	10	10
R5	F	31	10	21	9	8
E1	F	60	32	21	9.5	9.5
E2	M	34	12	21	8.5	8
E3	M	36	11	25	8	8
E4	F	44	19	15	10	10
E5	M	37	10	27	8	8.5
E6	M	32	10	22	9.5	9.5
C1	F	31	17	14	9.5	9.5
C2	M	26	10	16	8	9
C3	M	44	12	32	8	8
C4	F	37	10	27	8	8
C5	F	35	10	25	8	9
C6	F	45	20	25	9	8
Mean (sd)		40.3 (9.40)	14.8 (6.46)	23.8 (5.39)	8.85 (.75)	8.88 (.78)

Key: M/F=sex; LOR= length of residence in Norway; AGE=age at the time of test; ASSN=age started speaking Norwegian; NOC=self-rated Norwegian oral comprehension; NOP= self-rated Norwegian oral production.

All the informants reported that they had learned the Bokmål variety of Norwegian in school, and therefore spoke Norwegian following Bokmål grammar and pronunciation. Some of the informants noted that although they did not rate their oral comprehension as 10, they did not have trouble understanding people any more than other Norwegians do: even native speakers of Norwegian find it difficult understanding different Norwegian dialects. That is very true. There is no Standard spoken Norwegian, but a number of different varieties (the so-called dialects) spoken in Norway. Norwegian dialects can be so different that speakers of the two different dialects have difficulty understanding each other. The language situation in Norway may well influence the L2 speakers' rating of their oral comprehension of Norwegian. Nevertheless their reported levels of Norwegian oral comprehension are very high, as can be observed from Table 5.1 above.

²³ According to place and manner of articulation, Norwegian /l/ can be described as lateral approximant alveolar; Norwegian /r/ as trill alveolar (Askedal 1994:222).

It should be noted that Norwegian, strictly speaking, was not the informants' *second* language. As we can see from Table 5.2, almost all the Romance and the Chinese informants had learned some English before they started to learn Norwegian; two of the English informants had learned some French before starting Norwegian. So Norwegian had literally been their third or even the fourth language. In order to rule out the possibility that certain formal features were acquired before the critical period, the L2 informants were selected only when they had not acquired during their childhood a second language whose D and α roots carry an uninterpretable feature that their L1s lack (cf. inclusion criteria). Also note that except for R2, who reported that she was equally proficient in English and Norwegian, all the other informants rated their Norwegian proficiency level higher than that of other 'foreign' languages they had learned.²⁴

Table 5.2. L2 Informants' foreign language configurations and age starting the language(s)

ID	L2 (AGE)	L3(AGE)	L4(AGE)
R1	English (11)	Norwegian (29)	French (36)
R2	German (8)	English (8)	Norwegian (24)
R3	English (11)	Norwegian (32)	n/a
R4	English (12)	Norwegian (28)	n/a
R5	English (9)	Norwegian (21)	n/a
E1	French (14)	Norwegian (21)	German (24)
E2	French (14)	Norwegian (21)	n/a
E3	Norwegian (25)	n/a	n/a
E4	Norwegian (15)	n/a	n/a
E5	Norwegian (27)	n/a	n/a
E6	Norwegian (22)	n/a	n/a
C1	Norwegian (14)	English (14)	n/a
C2	English (13)	Norwegian (16)	n/a
C3	English (14)	Norwegian (32)	n/a
C4	English (14)	Norwegian (27)	n/a
C5	English (14)	Norwegian (25)	n/a
C6	English (14)	Norwegian (25)	n/a

Finally a brief note must be made on the identification of endstate L2 learners. Both Long (2003) and White (2003) suggest that the most satisfactory method of identifying an endstate grammar is by means of longitudinal data. If no change is found over time in the relevant aspect of grammar, this would suggest that a steady state has been achieved. Lardiere's subject, Patty, has been considered as the one of the best exemplars of endstate L2 acquirers. However, recently Lardiere (2006b) reported that Patty's grammar actually was still undergoing development, despite her long immersion in L2

²⁴ One may wonder whether all the previously acquired languages have possible transfer effects on the acquisition of the subsequent languages, as studies on third language acquisition (e.g. Leung 1998) seem to suggest. It should be pointed out, however, the facilitative effects were only observed in the initial state of acquisition. The effects are presumably less obvious in the end state, particular in the case of the informants in this study, where the Norwegian is their dominant foreign language.

English. Table 5.3 shows that Patty's plural marking had been improved by over 11% between her second recording (after ten years in the USA) and her third recording (nine years later).

Table 5.3. Patty's plural marking in obligatory contexts

	Suppliance/contexts	%
Recording 1	2/23	08.70
2	24/51	47.06
3	14/24	58.33

(source: Lardiere 2006b)

Despite the fact that Patty's grammar was still improving, Lardiere (personal communication 2006) still considers her as an endstate L2 speaker, on the basis that the percentage of improvement between the second and the third recording is much smaller compared with the one between the first and the second recordings. Patty's case indicates that absolute 'no change' in the interlanguage grammar is only an idealization—endstate interlanguage grammar may well undergo further (but not dramatic) development over time. In addition, gathering longitudinal data is not always possible for practical reasons, as pointed out by White (2003a). So the 'less satisfactory criteria' were adopted in this study, acknowledging the potential weakness of this methodology. Nevertheless, I assume that my informants had reached their ultimate attainment, based on their length of immersion in the L2 Norwegian, their self-rated (very high level of) oral comprehension and oral production of Norwegian. They lived and worked in a nearly exclusively L2 environment, and their exposure to the L2 Norwegian had been of sufficient duration so that we do not anticipate that their interlanguage grammars should undergo dramatic further development over time. In this aspect, my informants were comparable to Lardiere's subject, Patty. Still I acknowledge that "Proficiency is no substitution for longitudinal data" (Long 2003), so longitudinal data will be collected in the future research to determine whether the L2 informants are indeed at their end state.

5.2.2 Norwegian Proficiency Test (NPT)

Although the informants' self-rated oral comprehension and production of Norwegian reflected that they were communicatively highly proficient in L2 Norwegian, there is still a need for an independent and more objective measure of their command of Norwegian in various linguistic domains, including vocabulary, sentences, and discourse. Hence a standard test which can measure L2 learners' knowledge of Norwegian is needed. However, such a test is not readily available. So I designed one on my own, based on NTNU's Level-3 (the advanced level) final examinations for the Norwegian language course.²⁵ The proficiency test was in the form of a Close test, where 50 words were taken away from a 443-word long passage (see Appendix 1 for the Close test). The informants were asked to fill in one word for each blank. The

²⁵ Bergenstest 'Bergen test' is a standard test of Norwegian as a foreign language, which consists of a 6-hour written test and an oral test. Obviously, it is too long to be suitable for the informants in my study. Some of the informants had taken Bergenstest, but they were asked to take the proficiency test as well in order to provide a comparable measure with the others.

words to be filled in involved a wide range of word classes, such as nouns, adjectives, verbs, conjunctions, prepositions, adverbs, determiners. In addition to targeting DP internal agreement, it was also aimed at testing other aspects of grammar, as well as vocabulary and reading comprehension. No time limit was set for this test, but most of the informants completed it within 30 minutes. The informants' responses were considered as correct answers only if they were grammatically correct and suitable for the context. For many blanks, there could be more than one suitable word. The informants' answers were checked by the author. In case there was any uncertainty about the eligibility of the informants' answers, the author would consult a native speaker, who is an experienced teacher of Norwegian, and let her judge the appropriateness of those words. The informants' scores for the Norwegian proficiency test are listed in Table 5.4 below:

Table 5.4. Informants' NPT scores (maximal points: 50)

ID	R1	R2	R3	R4	R5	E1	E2	E3	E4	E5	E6	C1	C2	C3	C4	C5	C6
Score	48	47	48	49	48	50	47	48	50	47	49	48	48	46	47	47	39

The results indicated that the informants were indeed highly proficient in Norwegian, except for one, i.e. C6. C6 made 11 errors out of 50 blanks, indicating that she might lack metalinguistic knowledge of Norwegian grammar, though she had been using Norwegian in her everyday life for over 20 years. Due to her poor performance in the Norwegian proficiency test, she was excluded from subsequent experiments. As a result, 16 L2 informants (8 males, 8 females) were retained for the main experiments.

The remaining 16 informants did not differ much in scores on the NPT. The number of errors they made ranged from 0 to 4, suggesting that they had fairly good knowledge of the Norwegian grammar. Among the three groups, the English informants scored the highest, with 1.5 errors on average; the Romance group ranked in the middle, with 2 errors on average; the Chinese informants the lowest, with 2.8 errors on average. But one-way ANOVA ($F(2, 13) = 2.108$, $p = .161$) and post-hoc tests did not yield any significant differences within any two groups [in all cases, $p > .05$]. So we can assume these L2 informants were at the same proficiency level in terms of their L2 Norwegian.

5.2.3 L1 participants

The native control group was made up of fourteen native speakers of Norwegian (6 males, 8 females). They were generally age-matched with the L2 participants. Their mean age at the time of testing was 36.4 (sd = 10.6), ranging from 25 to 64. Most of the L1 informants were undergraduate or graduate students studying at the Norwegian University of Science and Technology (NTNU). A small number of informants were researchers or secretaries from the same university. Phoneticians and syntacticians were excluded. Hence the L1 informants were presumably comparable with the L2 informants with respect to intellectual capacities. The native informants came from different parts of Norway and spoke a variety of Norwegian dialects. But all had good

knowledge of Bokmål variety of Norwegian, because they had learned Bokmål in school.²⁶ No informants had any hearing impairments.

5.3 Experiment 1: Gender Assignment

5.3.1 Introduction

The aim of this experiment was threefold. First I wished to determine whether the L2 participants can assign the correct gender to the nouns to be used in experiments 2 and 3 (see below). A prerequisite for determining whether the informants can produce correct gender agreement in the on-line DP production task (see below in the present chapter) and whether they can perceive gender agreement cues in the auditory naming task (which will be reported in Chapter 6) is to know whether they can associate correct gender with the nouns. In this sense, the gender assignment task is an indispensable part for the subsequent experiments. Further, I wish to examine whether there are any differences among the three different groups in assigning gender to the Norwegian nouns. As has been discussed in the previous chapter, the L2 participants from the Romance group have gender in their L1s, but the gender system in their L1s is not congruent with the Norwegian gender system. English and Chinese do not have grammatical gender at all. So the question we ask is: will the presence or absence of a gender system in L1s results in any differences in the L2 participants' performance in the gender assignment task? Particularly, will the presence of a gender system in Italian/Spanish put the Romance group in an advantageous position over the other two groups? Lastly, I would like to examine whether the L2 informants are consistent in marking (D-N) gender agreement by comparing the results from the gender assignment task and the on-line DP production task.

5.3.2 Method

a. *Participants*: Only the three L2 experimental groups, namely, the Romance group (R), the English group (E) and the Chinese group (C), participated in this experiment. The Romance group was made up of 5 Italian/Spanish-speaking learners; the English group consisted of 6 English-speaking learners; the Chinese group was made up of 5 Mandarin-speaking learners. See Section 2 of the present chapter for detailed background information about the L2 participants. The native informants did not take part in the test, as it has been generally believed that native speakers have no problem acquiring the gender system of their own language.

b. *Materials and procedure*: Sixty-four Norwegian nouns were chosen for this study (see Appendix 2 for a complete list). They were composed of 20 nouns that were used in the indefinite singular conditions of the on-line DP production task (see below), 36 nouns from the part of the auditory naming task which targets gender agreement (see Chapter 6 for more details), and 8 nouns that were not included in either task. Among the 64 nouns, 34 were of neuter gender, 30 of common gender. They all belong to the

²⁶ The native informants were taught *Nynorsk* 'Neo-Norwegian' as well. But Bokmål had been learned as their primary language, Nynorsk as their secondary language.

top or upper-intermediate quarter of the Norwegian nominal vocabulary, based on the *Oslo Corpus of Tagged Norwegian Text*).

The task involved pencil and paper work. The informants were tested individually. They were given the list of nouns, each being preceded by both *et* (neuter indefinite determiner) and *en* (common indefinite determiner). They were asked to circle the correct gender for each noun. In case there were words they did not know, they were instructed to underline the nouns, and not to assign any gender to them. The informants followed their own pace doing this task; no time limit was set. This task was implemented before the on-line DP production task.

5.3.3 Results

All nouns were assigned a gender by all the informants, indicating that the informants knew the meanings of the nouns in the gender assignment task. Individual informants' correct percentages for neuter gender nouns (referred to as *et* items) and common gender nouns (referred to as *en* items) are listed in Table 5.5.

Table 5.5. Individual informants' accuracy on *et* items and *en* items

	<i>et</i> items (n=34)		<i>en</i> items (n=30)	
R1	27/34	79.4%	29/30	97%
R2	26/34	76.5%	30/30	100%
R3	32/34	94.1%	30/30	100%
R4	32/34	94.1%	29/30	97%
R5	33/34	97.5%	29/30	97%
E1	33/34	97.5%	30/30	100%
E2	29/34	85.3%	25/30	83%
E3	30/34	88.2%	26/30	86.7%
E4	34/34	100%	30/30	100%
E5	30/34	88.2%	25/30	83%
E6	32/34	94.1%	28/30	93.3%
C1	24/34	70.6%	30/30	100%
C2	32/34	94.1%	30/30	100%
C3	28/34	82.4%	28/30	93.3%
C4	30/34	88.2%	27/30	90%
C5	29/34	85.3%	27/30	90%

We can observe that R1 and R2 had relatively poor performance on *et* items. Both of them had an accuracy rate below 80%. In contrast, they were very accurate on *en* items, with an accuracy rate above 97%. This suggests that R1 and R2 took the more frequent *en* as the default gender. In case they were not sure about the gender of a certain noun, they would resort to the default gender strategy in assigning gender to that noun, resulting in the seemingly high accuracy rate on *en* items. The other three informants in the Romance group were very accurate both on *en* items and on *et* items, with less than 2 errors in each category. Their performance in gender assignment thus was target-like.

Three of the English-speaking informants, namely, E1, E4, and E6, were also target-like in their use of articles, almost assigning correct gender for every given noun. The other three English-speaking informants also had relatively high degree of accuracy (83-

86.7% for *en* items; 85.3-88.2% for *et* items), but nevertheless made a few errors on both *en* and *et* items. Surprisingly four informants displayed a similar accuracy profile: for each of them, the accuracy rates on *et* and *en* items were quite close. It seems that neither of the article pair was overgeneralized by these three English informants. However, it is not clear on what basis they determined the nouns' gender.

Turning to the Chinese group, we find that one informant, i.e. C2, was almost target-like in assigning gender to the Norwegian nouns. The other four informants all display the pattern of overgeneralizing *en*. The pattern was most clear in C1, who was 100% accurate on the *en* items in contrast with only 70.6% accuracy rate on *et* items. This is somewhat surprising, considering the fact that she had been exposed to Norwegian at a relatively young age (14), and had subsequently received all her education (from 6th grade in the primary school to Master degree) in Norwegian schools where the medium of instruction was Norwegian. It seems that her 17-year long exposure to Norwegian was no guarantee that she could be target-like in her use of articles.

All in all, there is individual variability in acquiring the grammatical gender system of the target language. The individual results suggest that some L2 learners can get close to target-like in gender marking (represented by R3, R4, R5, E1, E4, E6, and C2 in the present study), while the others seem to have persistent problem associating the right gender with Norwegian nouns (represented by R1, R2, E2, E3, E5, C1, C3, C4, and C5 in the present study).

After analyzing the performance of individual informants, we now look at the accuracy profile of the informants as a group. The group results reveal clearer tendencies that have been observed from the individual results. As we can see from Figure 5.1, a similar accuracy profile emerges for the Romance and the Chinese group, who were both more accurate on *en* items than on *et* items (98% vs. 88% for the Romance group; 95% vs. 84% for the Chinese group). The English group, however, were almost equally accurate on *en* and *et* items (93% vs. 92%). Hence, unlike the Romance and the Chinese group, it is not obvious whether the English group resorted to the default gender strategy in assigning gender to the Norwegian nouns. It turned out that, contrary to what we predicted, the English group behaved distinctly from the Romance and the Chinese groups. I do not have a good explanation for this. It could be due to the small sample size used in this study. It is not clear whether the accuracy profiles displayed by the five Romance-speaking learners, the six English-speaking learners, and the five Chinese-speaking learners, can be generalized to all endstate Romance, English, and Chinese speakers of Norwegian, respectively. Future study with larger samples in each group will reveal more reliable tendencies. Nonetheless, it seems safe to conclude that in the end state of second language acquisition, having a gender system (that is incongruent with Norwegian gender system) did not put the Romance group in an advantageous position over the English and the Chinese group, whose L1s do not have gender.

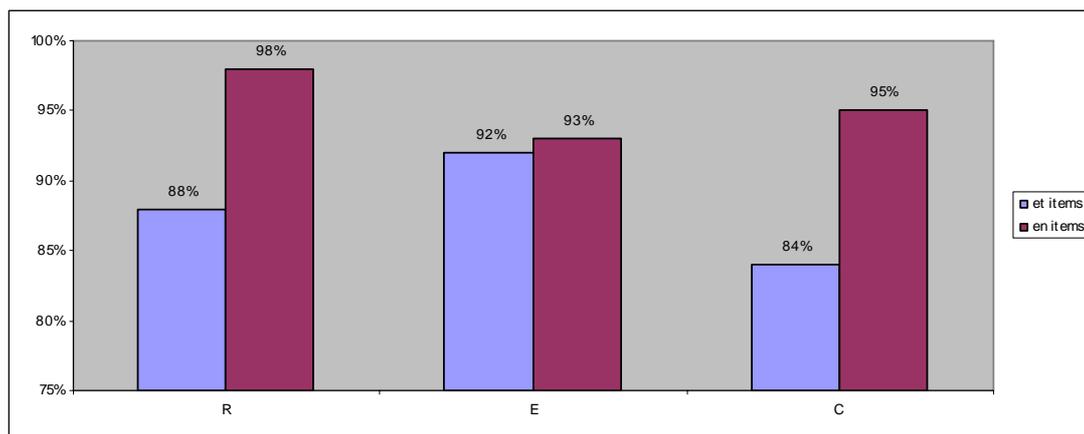


Figure 5.1. Each L2 group's accuracy on the gender assignment task

Despite the L2 learners' lack of total success, an average accuracy rate of 84% or above for each L2 group nevertheless suggests that the informants could assign gender to the nouns at the lexical level, though we need bear in mind that some individual informants performed less accurately on *et* items. We are still wondering whether the informants will be consistent in marking (D-N) gender agreement. Will the informants assign the same gender to the same noun in the on-line DP production task? The question can be answered after comparing genders assigned to the nouns that were shared by the gender assignment task and the DP production task. We will come back to this issue in the next section.

5.4 Experiment two: On-line DP production task

5.4.1 Introduction

The aim of the on-line DP production task was to elicit DPs containing adjectives. As we have known from the previous discussion, Norwegian has overt agreement between determiners, adjectives and nouns with respect to number, gender and definiteness. In addition, the inclusion of an adjective in indefinite DPs will not have any consequences on the pronominal determiners, but the inclusion of an adjective in definite DPs will result in the so-called double definiteness construction, which is considered as a clear manifestation of definiteness agreement in Norwegian. We are going to look at the Norwegian DP-internal agreement produced by the L2 informants, and determine whether the grammars of L2 speakers converge on the grammars of native speakers. The on-line mode of the experiment was selected in order to reduce the degree to which the participants have access to their metalinguistic knowledge. It is generally believed that on-line tasks are able to 'tap' L2 learners' implicit grammatical knowledge. The performance data from this experiment thus can, to a large extent, reflect the L2 speakers' grammatical representations of the target language.

5.4.2 Predictions

In light of the parametric differences between the target and source languages, summarized in Table 3.1, here repeated in Table 5.6, the main research question arises:

does interlanguage grammar of the endstate L2 speakers have the same properties as the native-speaker grammar? Or will the uninterpretable [*u*NUM], [*u*GEN], [*u*DEF] features be present in L2 speakers' interlanguage grammar depending on the status of the corresponding feature(s) in their L1s?

Table 5.6. Parameterized uninterpretable features in the L2 and the L1s

	D			α		
	[<i>u</i> NUM]	[<i>u</i> GEN]	[<i>u</i> DEF]	[<i>u</i> NUM]	[<i>u</i> GEN]	[<i>u</i> DEF]
Norwegian	+	+	+	+	+	+
Italian/Spanish	+	+	-	+	+	-
English	+	-	-	-	-	-
Chinese	-	-	-	-	-	-

Key: + Present in language; - absent in language

The current two competing SLA theoretical models, namely Full Transfer Full Access model (FTFA) and Failed Functional Features Hypothesis (FFFH) would have different predictions. The FTFA generally does not predict any differences between the three L2 groups. All the L2 participants are predicted to perform equally well on number, gender, and definiteness agreement, and they will acquire nativelike competence with respect to Norwegian DP internal agreement. Should there be L1-L2 divergence, however, the FTFA sees two possibilities. One is the Missing Surface inflection hypothesis (the MSIH), according to which errors in L2 speakers result from a mapping problem from abstract features to surface phonological forms. Specifically, following 'late lexical insertion' of Distributed Morphology, the MSIH argues that lexical items with unspecified features can replace those fully specified features on a syntactic node. Thus the realization of the inflected forms is predicted to be unidirectional, i.e. the unmarked, default forms occur in marked form contexts, but not vice versa. In the case at hand, lexical items on D and adjective are specified for features [\pm definiteness, \pm neuter \pm pl]. As has been discussed before, the features with positive values are marked forms, whereas those with negative values are defaults. That is to say, indefiniteness means absence of any definiteness features; common gender means absence of any gender feature; singular means absence of any number feature. Accordingly, the default form for D is *den* in definite contexts, *en* in indefinite contexts. The default form for adjectival inflection is less self-evident. According to Julien (2005: 46-47), of the three possible realizations of adjectival agreement /t/, /e/, zero 'ø', /t/ marks neuter singular, ø marks singular, /e/ is the unspecified realization of agreement. This is illustrated in (5.1, which is Julien's 2.31):

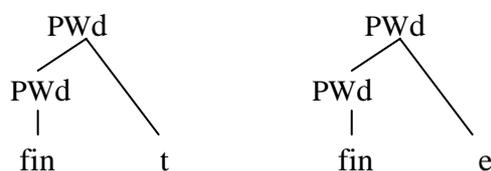
- (5.1) Realizations of the adjectival agreement in Norwegian
 a. [N SG] ↔ /t/ b. [SG] ↔ ø c. Agr ↔ /e/

Under the assumptions of the MSIH, we expect to find /e/ in /t/ and ø contexts, ø in /t/ contexts, but not vice versa. For the same reason, *den* is expected to be found in *det* and *de* contexts for the realizations of D.

Turning to the PTH, it is conceivable that Norwegian has the same adjunction structure with English for /t/ and /e/ inflection, as shown in (5.2) for *fin* 'fine'. But Chinese learners are expected to more easily adapt L1 prosodic structures to represent

/e/ than /t/, because it is possible for Mandarin syllables to end with /e/ (e.g. complementizer *de*, aspect marker *le*, sentence final particle *ne*) and other vowel sounds, but not with a consonant sound /t/.

(5.2) Prosodic structure of Norwegian adjectival inflections



In accordance with the PTH, we expect that Chinese learners' production of /t/ will be depressed in contrast to /e/, whereas English and Romance learners are not expected to make any errors with either inflection, because they should have no problem representing the prosodic structures of Norwegian. As to realizations of D elements, the PTH assumes that articles are organized higher in the prosodic tree, attaching directly to the phonological phrase (PPh), and predicts that such a structure will not be acquirable if it is not realized in the L1. The prediction, therefore, is that Chinese learners have persistent problems building up the appropriate prosodic representations for D elements in Norwegian, whereas the English and Romance learners should have no such problems. To sum up, the FTFA makes the following predictions:

- 1a) All learner groups are nativelike at the level of syntactic representations
- 1b) In case of errors, the MSIH predicts that realization of the inflected forms is unidirectional. That is, /e/ may occur in /t/ and \emptyset contexts, \emptyset in /t/ contexts, but not vice versa.
- 1c) The PTH predicts that English and Romance learners should be nativelike in realizations of D and suppliance of adjectival inflections; for Chinese learners, however, the suppliance of /t/ will be depressed in contrast with /e/, and realizations of D will be problematic.

On the hand, the FFFH, assuming that uninterpretable features not instantiated in L1 disappear beyond a critical period, predicts that the parametric differences between the target and source languages will result in:

- 2a) Definiteness agreement between determiners, adjectives and nouns will be problematic for all the three L2 groups, because the [μ DEF] feature is not present in their L1s.
- 2b) With respect to gender marking, Chinese and English speakers will behave differently from Romance speakers, who should perform much more accurately on gender agreement than both Chinese and English speakers.
- 2c) In terms of number agreement on determiners, Chinese speakers will behave distinctively from English and Romance speakers, both of whom perform more accurately than Chinese speakers.

5.4.3 Method

a. *Participants*: Participants in this experiment included the three experimental groups, namely, the Romance (R), the English (E), and the Chinese (C), and a control group, which consisted of 14 native speakers of Norwegian. Detailed information about the informants can be found in Section 2 of the present chapter.

b. *Materials*: Materials consisted of 60 short sentences (average length=5.2), all containing a simple nominal phrase, to be modified by an adjective. The nominal phrases appear either in sentence initial or sentence final positions. The nouns were everyday words which were supposed to be familiar to the L2 speakers. According to the gender, number and definiteness specifications of the nominal phrases, the test sentences were broken into six conditions, as illustrated in Table 5.7.

Table 5.7. Types of DPs and illustrations in the DP production task

	CONDITION	ILLUSTRATION
G1	INDEF. SG. COM.	De bodde i <i>en leilighet</i> . they lived in D-INDEF.SG.COM. apartment 'They lived in an apartment.'
G2	INDEF. SG. NEUT.	Jeg har <i>et problem</i> . I have D-INDEF.SG.NEUT. problem 'I have a problem.'
G3	INDEF. PL.	Han har mange <i>bøker</i> . he has many book-PL.INDEF. 'He has many books.'
G4	DEF. SG. NEUT.	<i>Klimaet</i> er ganske bra. climate-NEUT.SG.DEF. is quite good 'The climate is quite good.'
G5	DEF. SG. COM.	Jeg fikk låne <i>sykkelen hans</i> . I got borrow bicycle-COM.SG.DEF. his 'I was allowed to borrow his bicycle.'
G6	DEF. PL.	Alle reagerte på <i>boligprisene</i> . all reacted to house price-PL.DEF. 'All reacted to the house prices.'

There were 10 short sentences in each condition. The length of the sentence was controlled in order to lessen participants' memory load. No sentence exceeded 7 words, and the target sentences would not exceed 9 words. The sentences were audio-recorded by a native female speaker of Norwegian with standard Bokmål pronunciation at a natural speed.

Each nominal phrase in the short sentence stimuli was going to be modified by an adjective. The adjectives were not audio-recorded; they would be presented to the participants visually on a computer screen. Altogether 58 adjectives were chosen from the word frequency list of *the Oslo Corpus of Tagged Norwegian Text*. They all belong to the top quarter of the Norwegian adjective vocabulary in terms of written frequency. Two of the adjectives (*lang* 'long' and *annen* 'other') were repeated, making altogether 60 adjectives, one adjective for each sentence. The repeated adjectives were supposed to appear in different forms in the target DPs (i.e. *en lang replikk* vs. *det lange arbeidet*; *et*

annet tema vs. andre kulturer). Also care had been taken to make sure that all the adjectives which were supposed to modify singular indefinite nouns could have *-t* inflections. The modifying adjectives provided in the test were possible but not usual collocations with the noun. By so doing, I attempted to reduce the potential influence of familiarity in producing the Norwegian DP internal agreement. So if the participants produce DPs with no agreement errors, it is to a great extent an indication that they have internalized the rule, rather than simply being familiar with the phrases. The full set of short sentence stimuli, along with the modifying adjectives, is included in Appendix 4.

c. *Procedure*: The participants were tested individually in a phonetic lab. They were seated in front of a computer screen and a microphone. They first heard a short sentence presented to them over headphones and simultaneously saw on the computer screen a noun (in its bare form), which had occurred in the sentence, followed by an adjective (put in bracket). They were asked to make a new sentence, with the adjective modifying the noun, and to utter the new sentence loudly into the microphone. Taking the fourth sentence in Table 7 as an example, the informants would hear '*Klimaet er ganske bra*' and see on the computer screen 'klima (normal)', the expected response being '*Det normale klimaet er ganske bra.*' It was emphasized the new sentences should be made according to the standard grammar of Bokmål and that single definiteness of any kind should be considered ungrammatical.

The participants first went through three practice items with the experimenter still in the room. They were instructed that the nouns to be modified must be in the same form as they appeared in the stimuli (see Appendix 3 for the task instructions). The participants first practiced with the three examples by themselves. In case they had any misunderstanding of the task instructions, the experimenter would explain the task again, and let the participants listen to the sample responses. The participants would have to repeat the practice session until they totally understood the task instructions. For the main part of the experiment, participants were left alone. The test items were mixed and randomized, with an inter-stimulus interval set at 5 seconds. The order of presenting the stimuli was the same for all informants. The main part of the experiment took about 6 minutes.

d. *Transcription and coding system*: All the short sentences produced by the participants were transcribed by the author and double checked by a research assistant who specialized in Norwegian phonetics. Informants' failure to produce an utterance or failure to produce a DP as instructed was coded as Missed Items. The missed items were excluded from further analysis. The modified DPs produced by each of the participants were then coded according to the coding system below. In coding the DPs, N was taken as the reference point for agreement relations. Hence whether determiners or adjectives are correctly inflected for gender, number, and definiteness will depend on the inflections on the noun.

- CO=Correct DP — a well-formed DP with no agreement errors; CO also includes cases where the noun had been assigned the wrong gender (according to the gender assignment task), and both determiner and adjective agreed with the noun in (the wrong) gender. For example, *kirke* (common gender) 'church' was assigned a neuter gender by E2, and the corresponding DP he produced in the

on-line task was *et_{NEUT} greit_{NEUT} kirke*. This DP was coded as CO, because there supposed to be no agreement errors.

To pinpoint the informants' problematic areas, I divided the errors into the following six categories according to feature and domain combinations. Each category is illustrated below with error examples made by the L2 participants. Only the relevant features are annotated.

- **GenAdj**=Gender agreement errors in adjective— DPs in which the adjective had the wrong gender marking as related to the noun. (e.g. *et nær_{COM} land_{NEUT}; en aktivt_{NEUT} forfatter_{COM}*)
- **GenDet**=Gender agreement errors in determiner— DPs in which the determiner was incorrectly inflected for gender (e.g. *den_{COM} nye språket_{NEUT}; et_{NEUT} grei kirke_{COM}*)
- **NumAdj**=Number agreement errors in adjective— DPs in which the adjective was incorrectly inflected for number. (e.g. *fri_{SG} aviser_{PL}; en annen_{SG} kulturer_{PL}*)
- **NumDet**=Number agreement errors in determiner— DPs in which the determiner had the wrong number marking (*de_{PL} ville fjellet_{SG}; den_{SG} morsomme programene_{PL}*)
- **DefAdj**=Definiteness agreement errors in adjective— DPs in which the adjective was incorrectly inflected for definiteness (e.g. *det grønt_{INDEF} laget_{DEF}; et enorme_{DEF} ønske_{INDEF}*)
- **DefDet**=Definiteness agreement errors in determiner— DPs in which the determiner was incorrectly inflected for definiteness (e.g. *den_{DEF} enkel kode_{INDEF}; en_{INDEF} viktig byen_{DEF}*) and DPs in which the determiners were absent in contexts where they were obligatory (e.g. *ø sosiale symbolene_{DEF}*)

A few remarks on the error categories need to be made. First, the erroneous DPs produced by the participants may belong to more than one of the above error categories. For example, *spesielt demokratiet* falls into two categories, DefAdj and DefDet, because the adjective does not agree with the noun in terms of definiteness, and in the meanwhile, the prenominal determiner, which is obligatory in this context, is missing. Second, in many cases it is difficult to tease apart the number errors from the definiteness errors on adjectives, as *-e* is the plural marker on the adjective as well as the weak inflection. The solution here is that DPs were coded NumAdj only when the noun is in indefinite plural form, while the adjective is in the singular form (e.g. *frisk_{SG} nyheter_{PL}*). For erroneous DPs such as *de sosial symbolene*, *de positive ordene*, where the noun is in the definite plural form, while the adjective appears in the bare form, they were coded as definiteness agreement errors in adjective (DefAdj); in similar vein, DPs such as *en varme leilighet*, *et enorme ønsket*, where the noun is in indefinite singular form, whereas the adjective is wrongly inflected for *-e*, were also only coded as DefAdj, though Julien (2005) claims that the weak inflection is fully specified for gender, number and definiteness features. This coding method makes number and definiteness errors in adjective (NumAdj and DefAdj) mutually exclusive, but in the meanwhile it may result in a smaller number of NumAdj errors. We will come back to it in data analysis. Lastly, DPs of 'single definiteness' construction, such as *de_{DEF} lokale forkjeller_{INDEF}*, *den_{DEF} åpne oppførsel_{INDEF}*, were deemed definiteness agreement errors in

determiners (DefDet) errors, although ‘single definiteness’ is allowed in certain contexts in Norwegian (cf. Chapter 3). This is because in the experiment no context was provided that allowed single definiteness construction. In addition, it had been emphasized in the practice session that such a construction was not well-formed.

e. *Coding reliability*: DPs produced by each of the individual participants were initially coded and classified by the author. In order to check the reliability of coding, two samples (each sample consisting of all DPs produced by one participant) from each of the L2 groups were double checked by a researcher, who is native speaker of Norwegian and has a strong background in (Norwegian and general) syntax. In cases where the two raters disagreed about coding decisions, they would discuss the decisions until reaching agreement with respect to the error category assigned to each DP.

5.5 Results

5.5.1 Rate of missed items

Failure to produce an utterance or failure to produce a DP as instructed was deemed missed items. The average missing rate for the L1 group was 2.5% (21 per 840 trials, ranging from 0 to 7). One third of the missed items were cases where the adjectives were used as an adverbial instead of an attributive. The average missing rate for the L2 group was a 5.3% (51 per 960 trials, ranging from 0 to 15). Like the L1 group, a large proportion (19 out of 51) of missed items was due to failure to produce a DP according to instructions, rather than failure to produce an utterance. The small missing rates for L1 and L2 groups indicate that the length of the test sentences did not pose a particularly big challenge for the participants’ memory load, and the 5 second interstimulus interval allowed them reasonable time to produce a new sentence as instructed. The missed items were not subject to further analysis.

5.5.2 L1 group

As expected, no agreement errors were made by the native informants; all the modified DPs produced by the L1 participants were well-formed according to the standard grammatical rules of Bokmål Norwegian. It is noteworthy that no L1 participants produced any ‘single definiteness’ construction of any kind; in no cases did they omit the prenominal determiner or the postnominal determiner in contexts where ‘double definiteness’ was required.

5.5.3 L2 groups

a. *Overall accuracy rate*: I calculated the percentage of CO(rrect) DPs out of all the modified DPs produced by each of the individual participants (non-target-like DPs produced by the individual participants were summarized in Appendix 5). In order to give a more fine-grained analysis of individual performance, I also calculated the percentage of determiners and adjectives with correct inflections out of all the modified DPs produced by each individual participant. The percentage of correct DPs (referred to

as *absolute accuracy*), correct determiners (referred to as *accuracy on D*), and correct adjectives (referred to as *accuracy on α*) are presented in Table 5.8 below.

Table 5.8. Absolute accuracy, accuracy on D and α for each L2 participant

	Absolute accuracy (%)		Accuracy on D (%)		Accuracy on α (%)	
R1	53/59	89.8	56/59	94.9	56/59	94.9
R2	59/60	98.3	59/60	98.3	60/60	100
R3	29/52	55.8	32/52	61.5	37/52	71.2
R4	54/60	90.0	56/60	93.3	56/60	93.3
R5	54/60	90.0	56/60	93.3	57/60	95.0
E1	58/59	98.3	58/59	98.3	59/59	100
E2	38/60	63.3	40/60	66.7	43/60	71.7
E3	27/50	54.0	32/50	64.0	26/50	52
E4	60/60	100	60/60	100	60/60	100
E5	20/45	44.4	22/45	48.9	31/45	68.9
E6	59/59	100	59/59	100	59/59	100
C1	27/57	47.4	34/57	59.6	38/57	66.7
C2	37/58	63.8	47/58	81.0	45/58	77.6
C3	19/54	35.2	26/54	48.1	26/54	48.1
C4	25/59	42.4	36/59	61.0	42/59	71.1
C5	29/57	50.9	35/57	61.4	46/57	80.7

We can observe from the above table that there is some variability in the individual results within the Romance and the English groups. In the Romance group, R3 behaved radically differently from the others with his poor performance (an absolute accuracy of 51.9% as against above 89.8% for others); in the English group, the performance of E1, E4 and E6 contrasted sharply with that of E2, E3 and E5 in that the former three performed almost at ceiling (98.3-100%), whereas the latter three produced correct DPs around chance level (44.4%-63.3%). The Chinese group, however, behaved quite uniformly, producing correct DPs more or less at chance level (35.2-63.8%).

Summing up the individual results, it appears that the task divided the L2 participants into two groups, as can be observed from Figure 5.2. One group (including 4 informants from the Romance group and 3 informants from the English group) had nativelike performance, with correct percentages of 89.8% or above; while the rest of the nine informants (including one informant from the Romance group, 3 informants from the English group, and 5 informants from the Chinese group) produced correct DP internal agreement more or less at chance level. A chi-square test confirmed this observation: a significant difference was found between R1, who has the lowest correct percentage among the nativelike performers, and C2, who had the highest percentage among the non-nativelike performers ($\chi^2 = 10.456$, $df = 1$, $p = .001$).

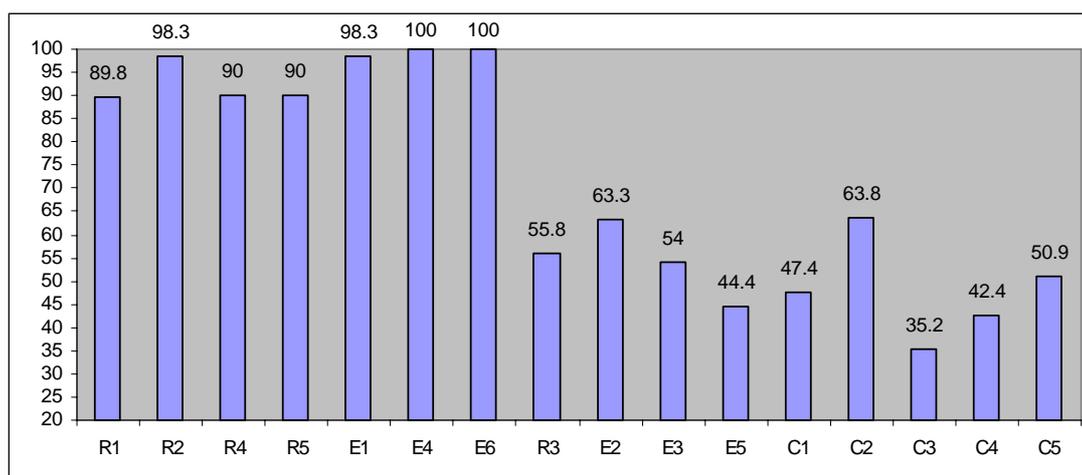


Figure 5.2. Accuracy rate of the DPs produced by the each L2 informant

At this juncture, one may question the validity of setting the criterion of 89.8% of accuracy as constituting nativelike performance. Indeed different (and often arbitrary) thresholds had been employed as criteria for reaching nativelike provision levels in the relevant literature (e.g. 60% in Vainikka and Young-Scholten 1994; 70% in Eubank and Grace 1998; but also see Hawkins and Franceschina 2003, where over 90% accuracy is still considered as non-native). In my case, the crucial point is that there is no statistically significant difference between R1 (whose accuracy rate is 89.9%) and E1 (whose accuracy rate is 98.3%) according to a chi-square test ($\chi^2 = 3.797$, $df = 1$, $p = .114$). Presumably no one will question the validity of setting the criterion of 98.3% of accuracy as constituting nativelike performance. Hence it is justified to label R1, R2, R4, R5, E1, E4, and E6 as nativelike performers.

If we examine the two groups' performance more closely, looking at their accuracy on D and on *a* respectively, we find that the individual L2 participants' accuracy rates on each domain were generally higher than their absolute accuracy rates, but still G2 participants are not on a par with the G1 participants in either domain, as can be observed from Figure 5.3.

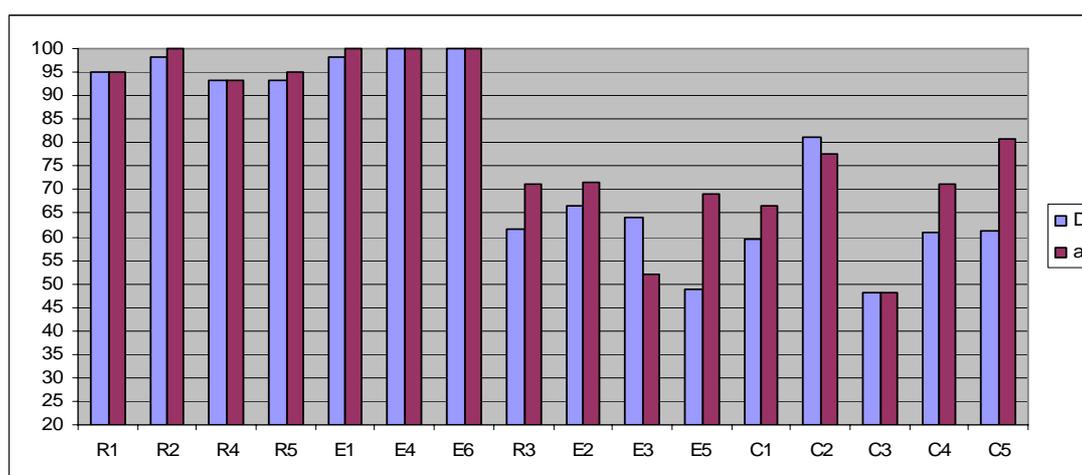


Figure 5.3. Accuracy on D and *a* for each L2 participant

Chi-square tests confirmed this observation: on the D domain, a significant difference was found between R5, who has the lowest correct percentage among the nativelike performers, and C2, who had the highest percentage among the non-nativelike performers ($x^2 = 4.020$, $df = 1$, $p = .045$); on the *a* domain, a significant difference was found between R1, who has the lowest correct percentage among the nativelike performers, and C5, who had the highest percentage among the non-nativelike performers ($x^2 = 4.177$, $df = 1$, $p = .041$). So again we found a clear contrast between the performances of the two groups, lending further support to our previous observations that G1 participants performed radically differently from G2 participants.

Based on the above observations, we rearrange the L2 participants according to their performance in the DP production task. Those who performed nativelike were assigned to group 1 (hereafter referred to as *G1*), and those who performed around the chance level to group 2 (hereafter referred to as *G2*).

b. *Realizations of adjectival inflections by G2*: We now focus on DPs produced by G2, which was made up of one informant from the Romance group, three informants from the English group, and all the informants from the Chinese group. We begin with a closer look at the realization of adjectival inflections by this group. In order to observe L1 effects, we further break G2 into three groups, according to the learners' L1s. Table 5.9 presents the percentage of erroneous adjectival inflections by each L2 group. Table 5.10 shows the Individual L2 group's suppliance of /t/, /e/, \emptyset in obligatory contexts.

Table 5.9. Each L2 group's errors on adjectives

Form used /Correct form	L1 Chinese		L1 English		L1 Romance	
\emptyset /e	28/170	16.47%	16/79	20.25%	8/30	26.67%
\emptyset /t	25/51	49.02%	15/34	44.12%	2/7	28.57%
e/t	2/51	3.92%	0/34	0%	0/7	0%
e/ \emptyset	5/61	8.02%	3/39	7.69%	0/7	0%
t/ \emptyset	4/61	6.56%	5/39	12.82%	1/10	10%
t/e	12/170	7.06%	4/79	5.06%	0/30	0%

Table 5.10. Each L2 group's suppliance of /t/, /e/, \emptyset in obligatory contexts

	L1 Chinese		L1 English		L1 Romance	
e	130/170	76.47%	59/79	74.68%	22/30	73.33%
t	24/51	47.06%	19/34	55.88%	5/7	71.42%
\emptyset	52/61	82.25%	31/39	79.49%	9/10	90%

Table 5.9 shows that the majority of errors are caused by the overuse of \emptyset in /e/ or /t/ contexts. By contrast, /e/ or /t/ is rarely found in non-obligatory contexts. This pattern of realizations of adjectival inflections is uniform across all the three learner groups. Though the observation that \emptyset occurs in /t/ contexts is in accordance with prediction 1b, the overuse of \emptyset in /e/ contexts is not compatible with it. According to the MSIH, /e/, which is the unspecified realization of agreement, is expected to be overused in /t/ and \emptyset contexts, contrary to what we have found.

We now examine whether the data support the PTH, which predicts that (only) Chinese learners' production of /t/ will be depressed in contrast to /e/ (cf. prediction 1c). The results from the Chinese group (cf. Table 5.10) appear to confirm this prediction: the suppliance rate of /e/ (76.47%) is indeed much higher than that of /t/ (47.06%). But it is important to note that the English group (though not the Romance group) shows exactly the same trend, i.e. the higher suppliance rate of /e/ as compared with /t/, which is not consistent with prediction 1c.

Errors on adjectives are expected under the FFFH. Results from Table 5.10 indicate that realizations of adjectival inflections across L2 groups generally do not reach native level. This suggests that learners have problems representing syntactic features in the target language. A fine-grained analysis of the errors will be provided later to see whether they are consistent with the predictions of the FFFH.

c. *Realizations of D by G2*: Due to a small number of indefinite contexts in the DP production task, we will not look into the realizations of indefinite Ds. We thus focus exclusively on realizations of definite Ds. Table 5.11 presents each L2 groups' suppliance rate of *den*, *det*, and *de* in obligatory contexts. Table 5.12 gives the percentage of errors on definite D in each L2 group.

Table 5.11. Each L2 group's suppliance of *den*, *det*, *de* in obligatory contexts

	L1 Chinese		L1 English		L1 Romance	
den	37/50	74%	7/17	41.18%	5/10	50%
det	15/43	34.88%	3/20	15%	0/7	0%
de	17/31	54.84%	4/14	28.57%	1/3	33.33%

Table 5.12. Each L2 group's errors on definite D

Form used /Correct form	L1 Chinese		L1 English		L1 Romance	
∅/den	8/50	16%	9/17	52.94%	5/10	50%
∅/det	5/43	11.63%	11/20	55%	6/7	85.71%
∅/de	5/31	16.13%	9/14	64.29%	2/3	66.67%
den/det	12/43	27.91%	5/20	25%	0/7	0%
den/de	8/31	25.81%	1/14	7.14%	0/3	0%
det/den	1/50	2%	0/17	0%	0/10	0%
det/de	1/31	3.23%	0/14	0%	0/3	0%
de/den	2/50	4%	0/17	0%	0/10	0%
de/det	9/43	20.93%	1/20	5%	1/7	14.29%

We find that L2 learners' realizations of definite Ds are not restricted to *den/det/de*, but also include a zero form ∅, as shown in (5.3). Table 5.12 shows that a majority of learners' errors on D can be attributed to the overuse of this zero form in *den/det/de* contexts. This means that determiners are omitted, instead of spelling out the default form *den*, contrary to what the MSIH predicted. Moreover, results from Table 5.11 indicate that the suppliance rate of *den/det/de* in obligatory contexts is low across L2 groups, which suggest that L2 learners as a whole have problems spelling out D in definite contexts. This observation is compatible with the FFFH, but is not expected

under the FTFA, or its proposals accounting for L2 errors, namely the MSIH and the PTH.

- (5.3) a. \emptyset dyre cd-en b. \emptyset grønne lag-et c. \emptyset mørke vintr-ene
 Target form: *den* dyre cd-en *det* grønne lag-et *de* mørke vintr-ene

d. *L1 effects*: To investigate the possible L1 effects, we now focus on the agreement errors made by the G2 participants. We wish to find out whether they made different types of errors depending on the presence and absence of uninterpretable features in their L1s, as the FFFH predicts. It should be recalled that the errors were divided into 6 categories according to feature and domain combinations: definiteness agreement errors in determiner (DefDet), number agreement errors in determiner (NumDet), gender agreement errors in determiner (GenDet), definiteness agreement errors in adjective (DefAdj), number agreement errors in adjective (NumAdj), gender agreement errors in adjective (GenAdj). Error rate (%) on each of the categories was calculated for each of the L2 groups within G2, and the results were presented in Figure 5.4.

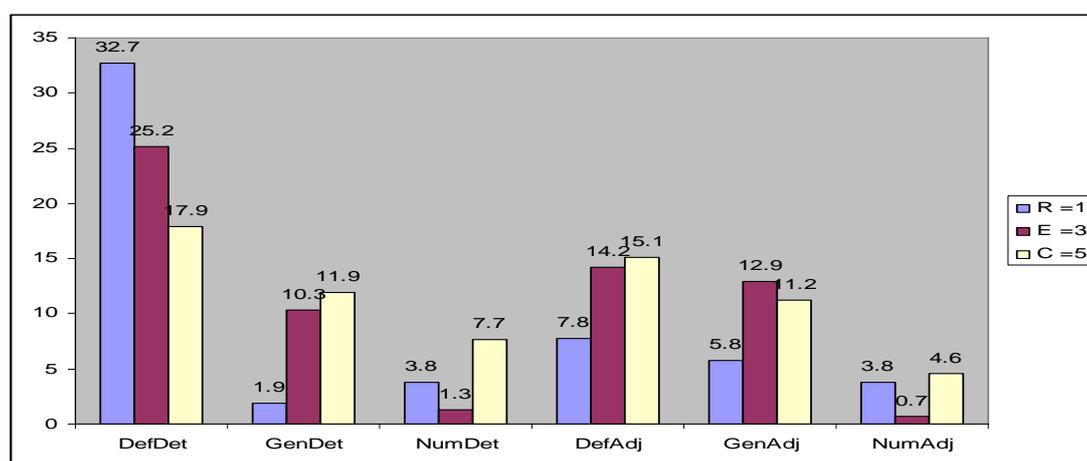


Figure 5.4. Individual L2 group' error rate (%) on each of the categories

As we can see from figure 5.4, all the three L2 groups made most errors in definiteness agreement on determiners (32.7% for the Romance, 25.2% for the English, and 17.9% for the Chinese group). Definiteness agreement errors on adjectives (DefAdj) ranked the next (7.8% for the Romance, 14.2% for the English, and 15.1% for the Chinese group). The definiteness agreement errors made by the G2 participants indicate that they all have problems accessing the uninterpretable definiteness feature on *a* and D in Norwegian. Since the [*u*DEF] feature on D and *a* have not been established, Agree is not invoked. Consequently the determiners and the adjectives are not spelt out target-like. Notice, however, the Romance-speaking informant made a relatively small number of DefAdj errors. This is probably because *a* root of his L1 carries the uninterpretable gender and number features, and these features are transferred to his L2 interlanguage grammar. The need to value the uninterpretable gender and number features on *a* root serves as triggers for spelling out the feature bundle on *a* in Norwegian, due to the impoverished nature of weak inflection. If this analysis is on the right track, the small number of DefAdj errors does not imply that he had acquired the [*u*DEF] feature on *a*.

With respect to gender agreement errors, the English and the Chinese groups' percentage for GenDet was over 5 times larger than that for the Romance group (10.3% and 11.9% vs. 1.9%); and for GenAdj about twice as high as that for the Romance group (12.9% and 11.2% vs. 5.8%). These differential behavioral patterns are closely related to properties of the D and *a* in the groups' respective L1s. While the Romance informants have established a [*u*GEN] feature both on D and on *a*, the English and the Chinese informants do not specify this feature on either head.

For number agreement errors on determiners (NumDet), the Chinese group made a relatively large number of errors, as compared against the English and the Romance groups (7.7% as against 1.7% and 3.8%). For NumAdj errors, however, the picture was less clear. It appears that all groups made few errors on this category (less than 5%). This is probably attributed to the method of coding NumAdj errors. As stated before, both the markers of definiteness and plurality on adjectives are *-e*, so it is not easy to tease apart the two types of errors. The solution adopted in this paper was to only code DPs with missing *-e* in INDEFINITE PLURAL context as NumAdj. It was largely due to this constraint that the occurrence of NumAdj errors was small. To put it another way, the smaller number of NumAdj errors resulted from fewer DPs on which this error can be made.

Ignoring the NumAdj category for the moment (as the number of errors in this category are small for all the three L2 groups), we now look at how the errors in each L2 group were distributed. We calculated the proportion of each error category out of a total number of errors made by each L2 group, and the results are shown in Table 5.13.

Table 5.13. Distribution of errors in each L2 group (%)

	DefDet	DefAdj	GenDet	GenAdj	NumDet
R	58.6	20.7	3.4	10.3	6.9
E	39.4	22.2	16.2	20.2	2
C	28	23.6	18.7	17.6	12.1

We see from Table 5.13 that the definiteness agreement errors (including DefDet and DefAdj) are predominant in the Romance group, making up approximately 80% of all errors; by contrast, number and gender agreement errors together are responsible for remaining 20% errors. For the English group, both definiteness agreement errors and gender agreement errors constitute a relatively large proportion (61.6% and 36.4%, respectively), but number agreement errors are rare (only 2%). Errors are fairly evenly distributed in each category in the Chinese group, indicating that the Chinese speakers are non-target-like with all these agreement features.

The error distribution for the three L2 groups clearly point to an L1 effect. That is, whether or not the participants made a certain type of errors is directly related to the presence or absence of the corresponding formal features in their L1s. It is those uninterpretable features which are not present in their L1s (either root D or root *α*) that were areas of persistent difficulty in these L2 speakers.

e. Consistency in gender marking: We now turn back to the issue of L2 speakers' consistency in marking D-N gender agreement. First we will determine whether the endstate L2 speakers are consistent in marking D-N gender agreement, by comparing the genders being assigned to the nouns shared by the off-line gender assignment task

and the on-line DP production task. The two tasks have 20 nouns in common, half in neuter gender, half in non-neuter (common) gender. As all the nouns were used in the singular indefinite situation of the on-line task, I first singled out all the singular indefinite DPs produced by each L2 participant, then compared the gender being assigned to the same nouns in the two tasks, and counted the number of the nouns whose gender alternated in the two tasks.²⁷ In addition, percentage for nouns whose genders were marked inconsistently was calculated for each of the L2 groups. The results are presented in Table 5.14.

Table 5.14. Rate of gender inconsistency at the individual and group level

ID	No. of sg. indef DPs	No. of Ns with alternate gender	Group percentage	
R1	18	4	16/95	16.8%
R2	20	5		
R3	18	2		
R4	19	2		
R5	20	3		
E1	19	1	16/116	13.8%
E2	19	3		
E3	18	4		
E4	20	0		
E5	20	7		
E6	20	1		
C1	19	5	17/97	17.5%
C2	20	0		
C3	20	2		
C4	20	6		
C5	18	4		

We find that L2 speakers indeed show persistent inconsistency in marking D-N gender agreement. All individual L2 participants, except for E4 and C2, exhibited a certain degree of inconsistency in the use of *en/et* with the same noun.²⁸ So it seems consistent gender marking is an exception rather than a rule even for endstate L2 speakers. If we look at the group percentage, we find a comparable degree of gender alternation among the three L2 groups: 16.8% for the Romance, 14.3% for the English, 17.5% for the Chinese (Chi-square tests did not yield any significant differences between any of the two groups [in all cases, $p > .05$]). Hence it appears that the Romance speakers, as a group, behaved similarly to the English and the Chinese speakers, whose L1s do not have a gender system. Here again, having a gender system did not help distinguish the Romance speakers from the other two groups of learners.

The group results, however, may conceal some important individual variation in selecting the indefinite article pair, if we consider the task-specific differences between the off-line gender assignment task and the on-line DP production task. There are at

²⁷ Some participants failed to produce the DPs as expected, resulting in a little individual variation in the number of indefinite singular DPs produced.

²⁸ But note C2 made some D-N gender agreement errors in the modified definite DPs he produced.

least two important differences between the two tasks. One difference is concerned with the accessibility of metalinguistic knowledge. In the off-line gender assignment task, no time limit was set; the informants could follow their own pace doing the task. In the on-line DP production task, however, the informants responded under great time pressure, hardly having any time to access metalinguistic knowledge. The other difference lies in the availability of the gender cue in the tasks. In the off-line gender assignment task, no gender cue was provided. The L2 participants could not rely on the phonological form of the nouns either, because there is no correlation between gender and the phonological form of the noun in Norwegian. In the on-line DP production task, however, the gender cue was provided by the indefinite articles (in the form of *en* and *et*) available in the auditory stimuli.

Given these characteristics of the on-line task, we wonder whether the L2 learners would ‘hear’ the gender cue available in the auditory stimuli and mark D-N gender agreement accordingly, or rather they would resort to the same strategy as they utilized in the off-line task?²⁹ If we find a participant who assigned a wrong gender to some nouns in the off-line task, but consistently supplied correct gender for the same nouns in the on-line task, this will be an indication that he/she hears the gender cue in the on-line task. Conversely, if correct gender had been assigned in the off-line task, but an incorrect gender was consistently selected for the same nouns in the on-line task, this will be an indication that the participant was not able to hear the available gender cue; instead he overgeneralized a gender form when performing the on-line task. In order to investigate the issue, we need to examine the gender alternation pattern of individual participants. This will constitute the first step examining the individual participants’ sensitivity to gender cue on D. It will also be of interest to compare G1 and G2 participants in order to determine whether the two groups differ in their ability to mark gender consistently and to hear the gender cue in the auditory stimuli. The nouns with alternate gender, their gender being assigned in the off-line task, and gender being produced in the on-line task for G1 participants and G2 participants are respectively summarized in Table 5.15 and Table 5.16 (E4 and C2 were not included, because they did not show inconsistency in gender marking).³⁰

Among the G1 participants, we notice that R1, R2, E1, and E6 wrongly assigned *en* to the noun (s) in the off-line task, but in the on-line task they selected the correct gender *et* for the same nouns. This is a clear indication that these participants could hear the gender cue in the auditory stimuli, and accordingly make use of it in marking gender agreement. R4, however, showed signs of overgeneralizing *en*, because she assigned the correct gender *et* to the nouns in the off-line task, but selected *en* for the same nouns in the on-line task. The pattern is mixed for R5, who in some cases seemed to be sensitive to the gender cue in the stimuli (reflected by his selection of correct gender for *våpen*), in other cases resorted to default gender strategy in the on-line task (reflected by his selection of the default gender *en* for *symbol* and *bygg*). Nevertheless, the number of nouns being supplied with the default gender was small, if we consider G1 as a whole. In general we can conclude that G1 participants were nativelike in terms of gender assignment and gender consistency except for R1 and R2, but crucially R1 and R2 were

²⁹ The phonetic forms of the two articles are: /en/ for *en*; /et/ for *et*, respectively. Since they end in different consonants, the phonetic difference between the two articles is quite large, and it is expected that the L2 speakers can hear the difference.

³⁰ A * symbol indicates that the gender is non-target-like.

able to ‘hear’ the gender cues in the auditory stimuli, and could correctly mark gender agreement. It is also noteworthy that for these participants, it is only neuter nouns that are subject to inconsistent gender-marking. The default gender *en* was overgeneralized to neuter contexts by these participants either in the gender assignment task or in the on-line task; no common gender nouns were marked inconsistently.

Table 5.15. Individual G1 participants’ gender alternation pattern

ID	Ns with alternate gender	Gender assigned (off-line)	Gender produced (on-line)
R1	tema	*en	et
	våpen	*en	et
	symbol	*en	et
	bygg	*en	et
R2	område	*en	et
	inntrykk	*en	et
	våpen	*en	et
	symbol	*en	et
	bygg	*en	et
R4	symbol	et	*en
	tema	et	*en
R5	symbol	et	*en
	bygg	et	*en
	våpen	*en	et
E1	symbol	*en	et
E6	ønske	*en	et

Turning to G2, we are struck by the large proportion of (26 out of 33) the nouns that were assigned a right gender in the off-line task, but were supplied with an incorrect gender in the on-line task, though the gender cue was available in the auditory stimuli of the on-line task. This can be an indication that these participants had difficulty ‘hearing’ the gender-marked determiners in the auditory stimuli. They did not have time to retrieve the appropriate determiner from their metalinguistic knowledge either, due to the time constraint in performing the on-line task. Consequently, one gender form was overgeneralized. A quick look at the gender alternation pattern showed that not all the participants overgeneralize *en* and that neither the English-speaking learners nor the Chinese-speaking learners behaved uniformly. We can observe that E2 overgeneralized *et* to non-neuter contexts in the on-line task, because he selected *et* for the two nouns to which a common gender *en* had been correctly assigned. E3 behaved alike, except in one case where he wrongly selected *en* for a neuter noun (i.e. *område*). E5 did the other way around: he selected *en* in the on-line task for the five neuter nouns to which *et* was properly assigned, indicating he took *en* as a default in the on-line task. He also appeared to properly supply a common gender to two common gender nouns to which *et* was wrongly assigned in the off-line task, but it is not clear whether it was because he noticed the gender cue in the stimuli or because he simply generalized the common gender to these nouns as well.

For the four Chinese-speaking informants, we can observe that there is also individual variability as to which of the article pair was overgeneralized. C1 and C4 seemed to overgeneralize *et* to the non-neuter contexts in the on-line task, because the nouns to which *en* was assigned had been supplied by the informants with a non-neuter gender *et* in the on-line task. This is surprising if we recall that it was *en* that had been overgeneralized by these two participants in the gender assignment task. C3 and C5, on the other hand, overgeneralized *en* in the on-line task, the same as they did in the gender-assignment task. R3 showed the same pattern as C3 and C5.

Table 5.16. Individual G2 participants' gender alternation pattern

ID	Ns with alternate Gender	Gender assigned (off-line)	Gender produced (on-line)
E2	våpen	*en	et
	sykkel	en	*et
	kultur	En	*et
E3	område	et	*en
	våpen	*en	et
	kultur	en	*et
	oppdragelse	En	*et
E5	område	et	*en
	bygg	et	*en
	sted	et	*en
	våpen	et	*en
	ønske	et	*en
	leilighet	*et	en
	kirke	*et	en
C1	kirke	en	*et
	oppdragelse	en	*et
	inntrykk	*en	et
	tema	*en	et
	symbol	Et	*en
C3	ønske	et	*en
	symbol	Et	*en
C4	våpen	*en	et
	kirke	en	*et
	leilighet	en	*et
	undersøkelse	en	*et
	kultur	en	*et
	forfatter	En	*et
C5	symbol	et	*en
	inntrykk	et	*en
	ønske	et	*en
	artikkel	En	*et
R3	symbol	et	*en
	ønske	Et	*en

In short, we found that G2 participants were generally unable to hear the gender cues provided in the indefinite articles in the on-line task. Instead they tended to overgeneralize one of the determiner pair. There was individual variability as to which determiner pair was overgeneralized: it was *en* for E5, R3, C3, and C5; *et* for E2, E3, C1, and C4. The surprising result is that an individual participant (namely, C1 and C4) could overgeneralize a different gender form, depending on the task—*en* in the off-line task, *et* in the on-line task. The underlying reason can be that these participants were aware of the fact that indefinite singular articles in Norwegian were gender-marked. But since [*u*GEN] had not been selected, the syntactic agreement mechanism had not been established. As a result, they could not consistently supply target-like gender-marked indefinite articles. Instead, they overgeneralized one of the articles. For some individual participants, both *en* and *et* were salient to them in the Norwegian input. Hence they selected different gender forms as default in different tasks.

Furthermore, both neuter gender nouns and non-neuter gender nouns are subject to inconsistent gender marking in G2 participants, in contrast to G1 participants, who only marked neuter gender nouns inconsistently.

5.6 Summary of findings and discussion

5.6.1 Summary of findings from experiments one and two

Examining the endstate L2 speakers' gender assignment and production of Norwegian DP internal agreement both at the individual level and at the group level revealed interesting and surprising results. The main findings from experiment one and experiment two are summarized below:

- Results from the gender assignment task showed that there is individual variability in acquiring the grammatical gender system of the target language at the lexical level. Some L2 speakers can get close to target-like in gender assignment; others have a persistent problem associating the right gender with Norwegian nouns; In addition, the Romance group and the Chinese group displayed a similar accuracy profile in the gender assignment task, overgeneralizing *en* to neuter contexts; the English-speaking learners did not show a tendency to overgeneralize either of the gender pair;
- Individual variability was also observed in performing the on-line DP production task. On the one hand it was found that four Italian/Spanish informants and three English informants performed nativelike (these learners were referred to as G1 participants), regardless of absence of uninterpretable definiteness feature in their L1s; on the other hand, there were one Spanish informant, three English informants, and all the Chinese informants, who produced DP internal agreement with an accuracy rate more or less at chance level (these learners were referred to as G2 participants);
- Neither adjectival inflections nor the definite Ds are realized as targetlike. Errors on adjectives are unidirectional: \emptyset is found in /t/ and /e/ contexts, whereas /e/ is rarely found in either /t/ or \emptyset contexts. For errors on Ds, the omission of D (or

the overuse of \emptyset form) is responsible for the majority of errors across learner groups.

- More fine-grained analysis of agreement errors made by the G2 participants showed an L1 effect— those uninterpretable features that are not present in the learners L1s (either root D or root α) were areas of persistent difficulty for these L2 speakers;
- Not all L2 speakers were consistent in marking gender agreement when comparing the results from the off-line gender assignment task with the on-line DP production task. There is a comparable degree of gender alternation among the three L2 groups: 16.8% for the Romance, 14.3% for the English, 17.5% for the Chinese;
- A closer examination of the individual gender alternation patterns revealed some individual variability. In general, G1 participants either could mark gender consistently or they were able to hear the gender cue provided in the indefinite articles available in the auditory stimuli; G2 participants, on the other hand, had difficulty hearing the gender-marked determiners when performing the on-line task. Instead they tended to overgeneralize one of the gender forms when it is necessary to supply an indefinite article on-line. Interestingly, there was individual variability as to which of the gender pair was overgeneralized. Six of them (namely, R3, R4, R5, E5, C3, and C5) overgeneralized *en*; four of them (namely, E2, E3, C1, and C4) overgeneralized *et*. Some learners (namely, C1 and C4) even showed optionality as to which gender was overgeneralized depending on the task: *en* in the off-line task, *et* in the on-line task.

5.6.2 Predictions revisited

The findings are not fully compatible with either the FTFA or the FFFH. According to the former model, a full set of features in the UG lexicon, interpretable or uninterpretable, continue to be available to endstate L2 speakers; according to the latter model, adult L2 speakers cannot represent those uninterpretable features not selected into the L1s. Both models are called into question by the existence of within-group variability in the Romance and the English groups. Within the Romance group, R3 behaved distinctly from other Romance-speaking informants with his poor performance in the on-line DP production task; within the English group, E1, E4, E6 had nativelike performance; by contrast, E2, E3, E5 produced target-like DP internal agreement at chance level.

The differential ultimate attainment within the Romance and the English groups seem to suggest that the parameterized uninterpretable features are selectively available to individual L2 learners. Some of the L2 learners are able to acquire the parametric uninterpretable features and have nativelike performance in terms of DP internal agreement, whereas others have persistent difficulty acquiring the uninterpretable features that are not realized in their L1s, and fail to produce target-like agreement associated with the parameterized uninterpretable features.

If this is the case, the following question naturally arises: what are the determining factors for the availability/unavailability of the formal features to individual adult L2 speakers? To put it in another way, what causes the difference in outcome between R1, R2, R4, R5 and R3, and between E1, E4, E6 and E2, E3, E5? I do not have a conclusive answer to this question. However, we know that it cannot be attributed to proficiency, as

all the L2 participants were assumed to be at the same proficiency level (based on their scores of the NPT, and their self-report oral and listening comprehension and production); or to intelligence, as all participants had received university level education or above, hence the same intellectual capacity was assumed; or age starting to learn Norwegian, as all were post-childhood L2 learners. The possible influence of another ‘foreign’ language had also been ruled out (cf. the inclusion criteria). The L2 participants differed in length of exposure to Norwegian, though, as reflected in their length of residence in Norway (LOR, ranging from 8 to 32). But a quick look at LOR of the L2 speakers, we find that three of the ‘nativelike’ L2 speakers, I1, I2, and E6 actually had relatively short LOR (11, 10, 12 years, respectively), as compared against some of those who were ‘non-native’ performers (e.g. R3 and C1, whose LOR were 20 and 17 years, respectively). So although no final conclusion can be drawn yet, length of exposure to Norwegian seems not to be a reliable predictor as to the failure or success of acquiring the parameterized uninterpretable features. The last factor I can think of is language aptitude. Although previous research (e.g. Dörnyei and Skehan 2003; DeKeyser 2000) shows that language aptitude may result in individual differences in second language learning, neither FFFH nor FTFA takes language aptitude into account. Both hypotheses assume that parameterized uninterpretable features are available or unavailable to ALL L2 learners, irrespective of their language aptitude. As the L2 participants’ aptitude had not been measured in the current study, I am not in the position to conclude whether aptitude plays a role in ultimate attainment by adult L2 learners. We have to leave this question into future research. But even if language aptitude IS a necessary condition for the ultimate success for L2 learners, as some researchers claim, we still have to seek to explain why it plays no role in the ultimate outcome in first language acquisition. Indeed if we consider the uniform outcome of first language acquisition, it is natural to conclude that language aptitude does not figure as a crucial factor in first language acquisition. It follows that we have to seek for other factors than language aptitude in order to account for the differential competence of L2 speakers from the same L1 backgrounds.

An alternative account, in line with the analysis of FFFH, is that the parameterized uninterpretable features are not available to adult L2 speakers, and that the nativelike performers’ apparent success with Norwegian DP internal agreement can be attributed to a misanalysis of Norwegian syntactic properties, making them compatible with the properties of the source languages. This is possible for at least Italian/Spanish-speaking informants. It should be recalled that in Italian/Spanish, the gender (as well as number) inflections on articles and adjectives generally follow the same pattern as the nouns. Two of the examples from Chapter 2, here repeated as (5.4), are suffice to illustrate the point:

- (5.4) a. i libri carini
 D-MASC.PL.DEF garden- MASC.PL small- MASC.PL
 ‘the nice books’
- b. le case carine
 D-F.PL.DEF house- F.PL nice- F.PL
 ‘the nice houses’

In the above examples, the inflections on ADJ and D have the same phonological shape with that of the N. The agreement pattern of the Norwegian double definiteness construction bears some resemblance with the Italian/Spanish pattern. Consider the phonetic forms of inflections on D and ADJ and N in the double definiteness construction, summarized in Table 5.17.

Table 5.17. Phonetic forms of inflections on D, ADJ, and N in the Norwegian double definiteness construction

FEATURE BUNDLES	D	ADJ	N
[SG][NEUT][DEF]	/de/	/-ə/	/-ə/
[SG][COM][DEF]	/den/	/-ə/	/-ən/
[PL][DEF]	/di:/	/-ə/	/-ənə/

The singular neuter definite determiner *det* has the phonetic form of /de/, ending with a central vowel; the corresponding suffixed determiner on the noun *-et* is reduced to an unstressed central vowel, a schwa.³¹ Similarly, in the common gender context, the phonetic form of the suffixed determiner is the reduced form of the prenominal determiner (/ə-ən/ vs. /en/). For the Italian/Spanish-speaking learners of Norwegian, one possibility is that the suffixed determiner on the noun is analyzed as gender and number markers, and the prenominal D is analyzed as the definite marker and is selected on the basis of the phonological form of the marker on the noun. This ‘misanalysis’ hypothesis, however, does not provide a satisfactory account for the success of some of the English-speaking informants, as English does not exhibit any DP internal agreement based on the phonological form of the noun.

Summing up, we provided two possible accounts for the within-group variability in the data. Under one account, the parameterized uninterpretable features may be selectively available to L2 learners. Under the other account, L2 learners may misanalyze the syntactic properties of the target language, resulting in their apparent success acquiring the uninterpretable features their L1s lack. Neither of the two accounts, however, is fully satisfactory in explaining the nature and extent of variability in the data. For the purpose of this research, suffice it to note that second language acquisition is a much more complex phenomenon than so far it has been conceived, and there can be many factors attributing to the success or failure of an individual L2 learner. The findings from this study also implicate the importance of analyzing data in terms of the performance of individual learners. In this study we found that L2 learners of the same group do not necessarily perform uniformly. Had we only considered the group results, many interesting findings would have escaped us. As White (2003) correctly points out, although group results can be quite informative, they can also be misleading, concealing properties of individual grammars.

Although neither model is fully compatible with the whole of the L2 data, the FFFH provides a better account for the non-target-like data than the FTFA. The FTFA resorts to the MSIH and the PTH in order to account for NS/NNS divergence. However, neither the MSIH nor the PTH made the right predictions for L2 errors on adjectives and definite Ds. Contrary to prediction 1b based on the MSIH, the unspecified realization of agreement on adjective /e/ is rarely found in /t/ or \emptyset contexts, and the default form *den* is

³¹ The *t* in both *det* and *-et* is silent.

not overused in *det* or *de* contexts. Prediction 1c based on the PTH is supported only by results from the Chinese learners, but is disconfirmed by results from the English and the Romance-speaking learners. For example, determiner omission found across L2 groups is not expected under the PTH. Moreover, neither the MSIH nor the PTH seems to be able to account for L2 learners' inconsistency in gender marking. All the above-mentioned errors are expected under the FFFH, according to which L2 learners are not able to represent novel uninterpretable features. A more fine-grained analysis of the non-target-like data indicates that definiteness agreement (on adjective and D) is problematic across L2 groups, which is in accordance with prediction 2a. Furthermore, errors with respect to gender and number marking point to L1 effects: L2 learners make gender/number agreement errors only when the feature is absent in their L1s. This finding is consistent with predictions 2b and 2c. Hence, the FFFH is supported by the data from the non-nativelike L2 performers.

5.6.3 Accounting for L1-L2 differences in production

Our ultimate goal is to find out whether or not post-critical L2 learners can ultimately achieve linguistic competence that is identical to that of a native speaker. The production data from the two experiments provide evidence that there are L1-L2 differences. At the production level, we found that post-critical endstate L2 learners differ from L1 speakers in the following two aspects:

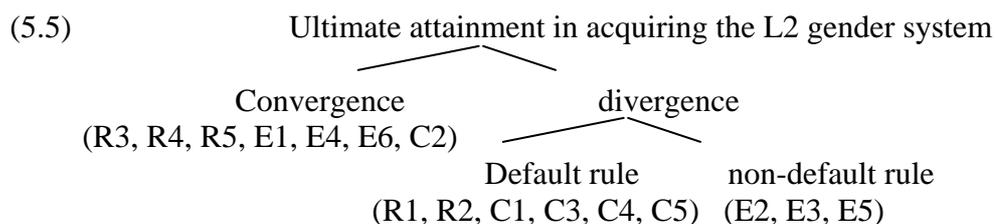
- learnability of the parameterized uninterpretable features
- learnability of the grammatical gender system at the lexical level
- Consistency in marking gender agreement

As to the learnability of the parameterized uninterpretable features, we found that in contrast to the ultimate success of L1 speakers, not all L2 learners can acquire the uninterpretable features their L1s lack beyond a critical period. What might account for this L1-L2 difference? It is generally assumed that the full set of features in the UG lexicon is available to children in the initial stage, and UG allows children to select language-specific features $[F_L]$ in response to the input they receive. Once the uninterpretable features are selected, they will drive syntactic computation, and the agreement features will spell out target-like.³² Features that are initially not selected are disregarded in the use of their L1. Among these features, it is the uninterpretable ones that cause learnability problems for learners to acquire a second language beyond childhood. It has been shown that the G2 participants did not acquire the uninterpretable features that were not realized in their L1s. However, the parameterized uninterpretable features can be available to some of the adult L2 speakers. The G1 participants, for example, have been demonstrated to have nativelike grammatical representations. In brief, results from the present suggest that the parameterized uninterpretable features are optionally available to adult L2 learners — some L2 learners can acquire these features, others cannot, though the reason underlying it remains unclear to us.

³² Certainly not all uninterpretable features are selected at the very beginning of language acquisition process. Selection of certain syntactic formal features may be delayed due to children's cognitive constraints or due to misanalysis of linguistic data. For Norwegian children, the prenominal determiner is usually absent in their early production of modified definite DPs (cf. Anderson 2004).

The second L1-L2 difference is concerned with learnability of the grammatical gender system at the lexical level. By ‘lexical level’, I mean the gender information that is inherent in the noun, which is supposedly interpretable on nominals. The lexical level needs to be distinguished from the syntactic level, where syntactic computation involving the matching of interpretable and uninterpretable gender feature is involved. The current study has revealed that there are cases where L2 learners (R1 and R2, for example) were target-like with gender at the syntactic level, making no DP internal agreement gender errors. But the two informants were not accurate with gender assignment at the lexical level, overgeneralizing *en* to neuter contexts. This is a clear indication that for adult L2 learners, the acquisition a gender system at the lexical level is not directly related to the acquisition gender agreement at the syntactic level.

Studies on first language acquisition of gender-marked articles (see for example, Carroll 1989 for French, Marinis 2003 for Greek) show that it is not an easy task for children to acquire the grammatical gender system at the lexical level, especially when the gender system is an ‘unnatural’ one, i.e. when the gender information on the noun is not directly related to the real-world gender. Like L2 learners, little children also go through a stage of overgeneralizing one of the gender forms (But also see Bohnacker 2004, where it has been shown that Swedish children do not undergo the overgeneralizing stage; gender marking is always targetlike for monolingual Swedish children). The difference between Children learning their L1 and adults acquiring an L2 lies in their ultimate attainment. The stage of overgeneralization does not last long for children; they soon become error-free in assigning gender to the nouns in their native language. Adult L2 learners, by contrast, often have persistent problems associating right gender with the nouns in the target language. As shown in this study, a number of the L2 informants kept using a default gender form even at the endstate, suggesting that this part of the grammar had ‘fossilized’: length of exposure to the target does not make a difference. This L1/L2 difference with respect to the ultimate attainment of gender systems lead many researchers to conclude that L1 and L2 learners may use different acquisition strategies: L2 learners might rely on a default gender strategy, while L1 learners may acquire the gender system by rote (see for example, Carroll, 1989; Bohnacker, 2004). An asymmetry gender system like that in Norwegian (with common gender twice as frequent as neuter gender) is likely to prompt L2 learners to overgeneralize common gender, and we see the residue of this default rule being used in the endstate L2 grammars (as in cases of R1, R2, C1, C3, C4, and C5). However, it has also been shown that there are cases where the endstate L2 grammars are divergent L1 grammars, but no default gender rule was used (represented by E2, E3, and E5). Moreover, there are L2 learners who could get close to target-like in gender marking, in which case the L2 grammars converge on L1 grammars (represented by R3, R4, R5, E1, E4, E6, and C2). The differential L2 ultimate attainment can be illustrated in (5.5) below:



A related L1-L2 difference is concerned with consistency in marking D-N gender agreement. While L1 speakers are always consistent in marking gender agreement, L2 speakers showed a considerable degree of gender alternation. When we compared the results from the off-line gender assignment task and the on-line DP production task, we found a comparable degree of gender alternation among the three L2 groups: 16.8% for the Romance, 14.3% for the English, 17.5% for the Chinese.

For L1 learners, the consistent gender marking can be attributed to the specification of a [*u*GEN] feature on D and the acquisition of the gender system at the lexical level. The uninterpretable gender feature drives the Agree operation, establishing a probe-goal relation between D and N, through which the [*u*GEN] feature on D assumes the value of the goal, i.e. the N. Consequently, D always agrees with N in terms of gender.

By logic, L2 learners' inconsistency in gender-marking can either be due to the absence of a [*u*GEN] feature on D, or due to the learners' inability to acquire the gender system at the lexical level, or both. For Chinese and English-speaking learners, inconsistency in gender-marking can be attributed to both factors, but the absence of a [*u*GEN] feature is presumably the main factor. When the [*u*GEN] feature is absent, agreement 'checking' mechanism is not available, hence D-N gender agreement is not guaranteed. For Romance-speaking learners, [*u*GEN] features are transferred from their L1s into their interlanguage grammar, and the checking mechanism is also in place. Inconsistency in gender-marking then can only be attributable to the problem of acquiring the gender system at the lexical level. As far as the present study is concerned, there can be task-specific reasons why gender inconsistency was induced. First, the on-line DP production task differs from the off-line gender assignment task with respect to the accessibility of metalinguistic knowledge. Even though L2 learners know the gender of the noun at the metalinguistic level, they may well fail to access to metalinguistic knowledge under time pressure, and resort to default gender strategy in their suppliance of articles. Second, gender cues were provided in indefinite determiners available in the auditory stimuli in the on-line task, but no gender information was available in the off-line task. So for those participants (namely, R1, R2) who could 'hear' the available gender cues, the gender marking inconsistency was caused by a discrepancy between the gender information in their metalinguistic knowledge and the gender cues provided in the auditory stimuli.

To conclude, it has been found that interlanguage grammars can be divergent from L1 grammars in many important aspects. What is most crucial for this study, however, is the finding that at least some post-critical L2 learners can ultimately achieve linguistic competence that is identical to that of a native speaker. Results from the on-line task suggest that G1 participants had developed nativelike grammatical representations with respect to Norwegian DP internal agreement, though G2 participants have fallen short of a nativelike grammatical knowledge in this respect. In the next chapter, we are going to determine whether end-state L2 speakers behave in a nativelike fashion when processing DP internal agreement violations. The issues of L1 transfer in L2 processing and the interrelation between L2 perception and production will also be addressed.

6 *L1 & L2 Perception*

6.1 Introduction

The production data from the on-line DP production task in the previous chapter have shown that some of the L2 participants (referred to as G1 participants) had nativelike performance; the remaining L2 participants (referred to as G2 participants) had problems producing target-like Norwegian DP internal agreement. The data suggest that the G1 participants were likely to have obtained nativelike linguistic knowledge, while the G2 participants fell short of nativelike grammatical competence, as evidenced by their respective performance in the DP production task. A fine-grained analysis of the agreement errors by G2 participants pointed to a strong L1 effect—the errors they made were restricted to those grammatical agreement properties that were not realized in their L1s. This means that the uninterpretable features which are present in L2 learners' L1s are also available in their L2 grammars.

The purpose of this chapter is to examine whether endstate adult L2 speakers have identical competence to that of native speakers at the level of processing. We are also interested in examining the interrelation between production and perception, that is, whether those participants who had nativelike competence at the level of production also have nativelike competence at the level of processing.

The experiment presented in this chapter was designed to investigate the on-line processing of Norwegian DPs by both native speakers of Norwegian and L2 speakers of Norwegian. The task is on-line in the sense that once an input element has fulfilled its function, it is discarded, in other words, it is no longer referenced. The experimental technique used in order to investigate L1 and L2 speakers' processing abilities is called *auditory naming* (alternatively called *cued shadowing*, or *spoken word repetition*, see Bates & Liu 1997), in which subjects are asked to listen to pairs of phrases and repeat a target word embedded in the carrier phrases. The cue-carrying item, usually appearing before the target word, is referred to as a *prime*. The time one takes to name a target word is called reaction time (RT, alternatively, repetition time, or naming time). RTs are measured in order to examine the effects of the prime on the recognition of a target. This technique has been used with great success to examine the degree to which L1 and L2 speakers are sensitive to local domain syntactic or semantic violations (see Bates and Liu 1997, for a review).

This chapter is structured as follows. In Section 6.2, I offer a brief review of the literature on L1 and L2 processing of local domain feature mismatches, arguing for a syntactic mechanism of the observed priming effects. This is followed by a proposal for a model of grammatical agreement processing, which draws on the agreement mechanism developed in recent versions of the Minimalist Program (MP). This model serves as the starting point for designing the auditory naming experiment, the details of which are presented in Section 6.3. Subsequently, in Section 6.4, I present results from the experiments, first for L1 participants, and then for L2 participants. In Section 6.5, the findings from the experiment are summarized and the observed L1/L2 processing differences are discussed in relation to L1 and L2 processing theories, particularly Shallow Structure Processing proposed by Clahsen and Felser (2006). Finally, a conclusion will be reached in Section 6.5.

6.2 L1 and L2 feature processing

Previous research on the effects of agreement cue on the recognition of noun targets, using auditory naming technique, has largely focused on gender-marking (among others, see Grosjean *et al.* (1994), for French; Bates *et al.* (1995, 1996) for Italian; van Berkum (1997) for Dutch; Akhutina *et al.* (1999) for Russian; Andonova *et al.* (2004) for Bulgarian). Researchers often used a gender concordant condition, discordant condition, and a neutral (without any gender information) condition. The neuter condition serves as the baseline against which both concordant and discordant conditions are compared. The effect is said to be *facilitative* if concordant cases are responded significantly faster relative to neutral cases; it is said to be *inhibitory* if discordant cases are responded significantly more slowly relative to neutral cases. Both facilitation and inhibition effects are referred to as indicating sensitivity to agreement cues on the prime. In general, it has been found that concordant gender-marking cases were responded to faster than both neutral and discordant gender-marking cases, suggesting that native speakers of gender-marked languages are sensitive to the gender cue when processing their native languages.

This gender marking effect has been explained as either due to a lexical module (e.g. Grosjean *et al.* 1994), or a syntactic module (e.g. van Berkum 1996), or a combination of a lexical module and a syntactic module (e.g. Bates *et al.* 1996; Grosjean *et al.* 1994, Guillelmon & Grosjean 2001). According to the lexical module approach, the influence of gender marking effect is internal to the narrow lexicon. The gender information on the prime is used to activate a noun set carrying the same gender feature, which facilitates the lexical search for the N target that has been activated in this set. The syntactic module approach, on the other hand, places the locus of gender marking effect on a post-lexical syntactic checking mechanism. This syntactic checking mechanism will see to it that grammatical agreement is respected. The violation of grammatical agreement will result in a “processing catastrophe” (Grosjean *et al.* 1994).

In contrast to large bulk of studies using auditory naming in L1 processing, there has been little empirical investigation using this technique in L2 research. To my knowledge, Guillelmon and Grosjean (2001) is the only one of this kind. In their study, they use an auditory naming task to examine how early English-French bilinguals (average age of onset of bilingualism is 5; 4), and late English-French bilinguals (average age of onset of bilingualism is 24; 8) react to grammatical gender violations

when processing French DPs. They find that the early bilinguals show clear gender priming effects, involving both facilitation and inhibition, whereas late bilinguals are totally insensitive to gender marking in perception.

This fact was interpreted by the author as related to the maturity hypothesis to second language acquisition. They argue that early bilinguals, like native speakers, can make use of the gender cue on D to activate lexical search and the syntactic checking mechanism, whereas certain processing mechanism in a second language will not be acquired beyond a certain age point. Hawkins and Franceschina (2003) reinterpret Guillelmon & Grosjean's findings. They argue that the early/late bilinguals' difference in processing reflects an underlying syntactic contrast between the two grammars, that is, an uninterpretable gender feature on D is available to early bilinguals, whereas it is not accessible to late bilinguals, due to the hypothesized constraints. But in an effort to explain the processing mechanism, they maintain that gender information influences lexical activation. They write:

The absence of such a feature means that in parsing, when late bilinguals have assigned *le*, *la* or *leur* to the category D, that is the end of the parse;³³ there are no consequential effects on the parsing of N. However, when monolinguals encounter *le* or *la* not only do they assign them to D, but that assignment calls up a [ugender] feature with either a [+fem] or [-fem] value, which activates the appropriate set of root Ns in the narrow lexicon, [+fem] or [-fem]. This speeds up lexical search for the N because only those Ns with appropriate gender features need to be accessed. If there is a clash between the feature of D and the feature of N, however, this will slow the search down, because the parser has initially activated the wrong set of Ns (Hawkins and Franceschina 2003: 22).

Hawkins and Franceschina claim that an uninterpretable gender feature is responsible for lexical activation when processing French, but it is not immediately clear under their account how come the uninterpretable gender feature on D, which is generally assumed to “enter the derivation unvalued” (Chomsky 2001) has been assigned “either a [+fem] or [-fem] value” at all. Moreover, if the L1/L2 processing differences lie in an underlying contrast in syntax between L1 and L2 grammars as they claim, it is reasonable to place the gender marking effect on a post-lexical, syntactic checking mechanism.

Independent evidence for the syntactic mechanism of feature processing is provided by experiments using the electroencephalogram (EEG) technique. This technique measures the electrical activity of the brain, by which event-related brain potentials (ERPs) can be obtained. Specifically, P600 effects are believed to be associated with syntactic processing difficulty. Researchers using this technique to investigate L1 processing of gender agreement violations found that native speakers indeed showed a P600 response to gender mismatch between determiners and the target nouns (see Hagoort and Brown 1999 for Dutch; Gunter *et al.* 2000 for German). Findings from the ERP studies thus further support the idea that a post-lexical, syntactic mechanism is involved in feature processing.

However, psycholinguistic studies usually stops at drawing a conclusion about the involvement of syntactic factors in processing, without probing into details as to how this syntactic mechanism works during language processing. To provide such syntactic details, one needs to draw on a syntactic framework that offers a precise mechanism of

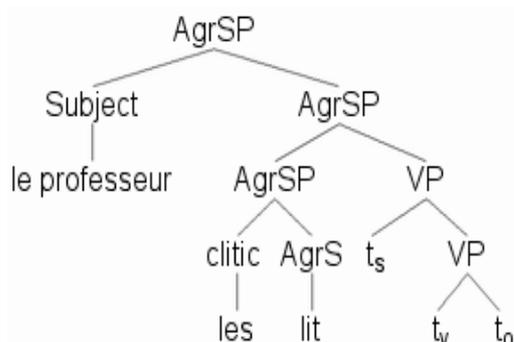
³³ *Le* is the singular masculine definite article in French, *la* is its feminine counterpart; *leur* (their) serves as neutral condition.

grammatical agreement. The study by Franck *et al.* (2006) is probably one of the first examples using linguistic theories as a tool to interpret findings from psycholinguistic experiments. They conducted a series of experiments examining the impact of structural relations on the occurrence of ‘attraction’ errors (that is, incorrect agreement with a word that is not the subject of the sentence). In one of their experiments, they used the test items like the following to elicit subject-verb number agreement errors in French:

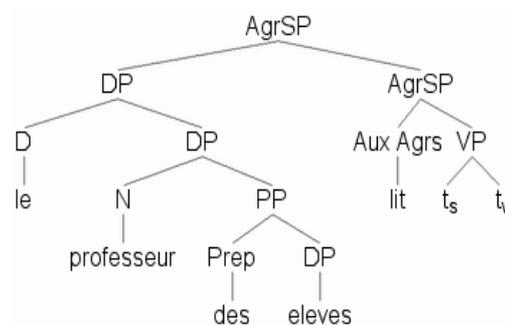
- (6.1) a. Le professeur les lit/*lisent
(The professor them reads/*read)
- b. Le professeur des élèves lit/*lisent
(The professor of the students reads/*read)

In (6.1a) an object clitic (*les* ‘them’) intervenes between the subject and the inflected verb; in (6.1b) a subject modifier (*élèves* ‘the students’) intervenes in terms of precedence between the head noun (*le professeur* ‘the professor’) and the verb. These two types of intervention are thought to be different in nature, as the hierarchical relations between the intervening constituent and the inflected verb in the syntactic representations are different. The syntactic representations of (6.1a) and (6.1b) assumed by the authors are shown in (6.2a) and (6.2b)

(6.2a)



(6.2b)



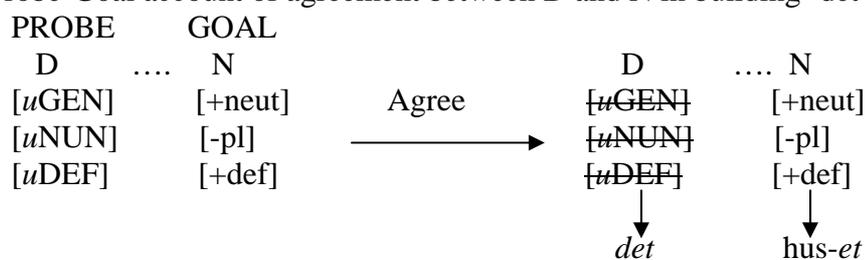
As can be seen from (6.2a), the subject c-commands the clitic, which in turn c-commands the inflected verb, so the clitic intervenes in terms of c-command; whereas in (6.2b) the subject modifier does not c-command the inflected verb; it only intervenes in terms of precedence. Franck *et al.* found that attraction in subject-verb agreement is significantly stronger when the intervening constituent is in a c-command relationship with AgrS (as in example 6.1a.) than when it is in a simple precedence relation to it (as in example 6.1b). The authors show that this finding can be explained based on the subject-verb agreement mechanism developed in the Minimalist framework. The main idea is that subject-verb number agreement involves the checking of the number features between ArgS and the subject, the two of which are in a c-command configuration. The authors note that “agreement errors arise precisely when the system deals with number specification, which is itself closely linked to agreement processes.”

Their study thus nicely illustrates the relevance and the benefit of combining the experimental approach of psycholinguistic and the refined analytic tools offered by formal syntax.

In the present study I wish to show that linking experimental psycholinguistics and syntactic theories can also help us better understand the syntactic mechanism involved processing of DP internal agreement violations. With this purpose in mind, I propose a model for grammatical agreement processing, which incorporates a Probe-Goal account of grammatical agreement developed in recent work of the Minimalist Program (MP) (Chomsky 1995, 1999, 2000). It is assumed that syntactic computation takes place during the syntactic formation of an utterance and during its parsing as well (Correa *et al.* 2005; Fong 2004; Phillips 2004). Applying the grammatical agreement mechanism developed in the MP to the target language of the present study, we can depict the structure building process of the Norwegian DP *det gamle huset* in the following way.³⁴ First the correct components for assembly are laid out. In this case, the components include D with uninterpretable gender ([*uGEN*]), number ([*uNUM*]), and definiteness ([*uDEF*]) features, adjective *gammel* with the same bundle of uninterpretable features (cf. Julien 2003, 2005), and *hus* with valued gender, number and definiteness features (for expository purposes, we refer to them as [+neut], [-pl], [+def], respectively). The assembly begins with a series of Merge operations, which results in D being in a c-command configuration with N.³⁵ The uninterpretable features on D render it a probe, searching in its c-command domain for a matching goal. It finds N, which carries the matching interpretable features. The operation Agree applies automatically as soon as the Probe-Goal relation between D and N is established, and through it, the unvalued features on D are deleted, and simultaneously the feature values on N are copied onto D. These newly gained features values on D spell out as *det*.

This process can be schematically represented in (6.3) below:

(6.3) Probe-Goal account of agreement between D and N in building ‘det gamle huset’



Where ‘....’ means c-command, a ~~strikethrough~~ means valuation and deletion of uninterpretable features, and a ↓ means ‘spell-out’.

In contrast to the bottom-up nature of this syntactic building process, parsing is “incremental and from left-to-right in nature” (Fong 2004). In this sense, parsing is decomposition of the phrase building process. Prior to parsing, lexical items are not

³⁴ For the purpose of this experiment, I only focus on the agreement between D and N. There is a Probe-Goal relation between α (with an Adjective Phrase in its Spec) and N as well (see the previous discussion in Chapter 4 or Julien 2005 for a detailed account), but this operation presumably has no effect on processing in my case, as an invariant adjective (i.e. *gamle*) has been used in all the test items of the present experiment.

³⁵ Two constituents X and Y are sisters if they are directly merged with each other. X c-commands its sister Y and any constituent Z that is contained within Y (see Chomsky 2000).

available. Due to this constraint, the assembly of phrase structure must proceed through elementary tree composition, rather than using the generative operations directly (Fong 2004). Elementary trees are “basically projections of functional and lexical heads” (Fong 2004), with (interpretable and uninterpretable) features specified. Accordingly, in parsing a Norwegian DP, an elementary tree of DP will be selected as soon as a D element is discovered. Once a DP is analyzed as such, the uninterpretable features on D will automatically drive the parsing process, through which a Probe-Goal relation between D and N will be established. If D and N have matching features as requested, the parsing will be efficient. On the other hand, if there is a single mismatch of features, attention will be directed to that feature mismatch, resulting in slowing down of processing. As we see, it is the uninterpretable features that drive syntactic computation and identify the probe and goal. The Probe-Goal relation between D and N explains why the information on D affects the processing of N.

Such a model offers an account of priming effects without the need of assuming a neutral condition, which was hit with much criticism and was conceived as “artificial and unrealistic starting point” (Bates *et al.* 1996). Moreover, it would predict that priming effects be found for all grammatical agreement features, not being restricted to gender-marking alone. If we assume that L2 processing of local domain grammatical agreement involves full syntactic computation, our agreement processing model would predict that L2 learners, especially highly proficient L2 learners, should react to agreement violations in a nativelike manner when their L1s and the L2 share similar agreement properties.

Assuming the agreement processing model proposed above, the finding from Guillelmon and Grosjean’s study can be interpreted this way: the [*u*GEN] feature is available to the early bilinguals, so like native speakers of French, they have specified a [*u*GEN] feature on D, which forces the Agree operation to take place for the purpose of feature valuation. The matching or mismatching of features between the probe and goal will subsequently facilitate or inhibit processing, resulting in early bilinguals’ sensitivity to the gender cue on D. Conversely, the late bilinguals have not specified a [*u*GEN] feature on D in their interlanguage grammar of French, as Hawkins and Franceschina claimed. So a Probe-Goal relation between D and N with respect to gender will not be established. For this reason the gender information on D presumably has no effects on the processing of N. Accordingly, the late bilinguals are insensitive to gender agreement violations in French.

For the moment I suggest that the lack of [*u*GEN] feature in the L2 grammar could be a possible source of the late bilinguals’ insensitivity to gender violations in L2 French. As Guillelmon and Grosjean did not include an L2 group whose native language also have gender agreement properties, it is impossible to determine whether L2 speakers will show nativelike processing when L1 and L2 are syntactically similar. In addition, there is no relevant data in their study showing whether the late bilinguals have nativelike performance with respect to gender agreement in French. So there is no way to tell how the late bilinguals’ perception is related to their production. In order to assess the possible influence of factors such as grammatical competence and L1 transfer on L2 processing, the present study looks at three groups of L2 learners, Romance (similar to L2 Norwegian in terms of number and gender agreement), English (similar to L2 Norwegian in terms of number agreement), and Chinese (different from L2 Norwegian in terms of all the agreement properties). Definiteness agreement is unique

to L2 Norwegian. These L2 groups from typologically different language backgrounds will enable us to probe the L1 transfer effects on L2 processing. Moreover, these L2 speakers' linguistic knowledge regarding Norwegian DP internal agreement had been tested using an on-line DP production task. Based on their performance in that task, they were broken into two groups: G1, which was made up of 'nativelike performers', and G2, made up of 'non-nativelike performers'. A comparison of the processing performances between these two groups will enable us to determine the interrelation between the L2 production and perception. I should be made it clear however, I am not claiming that the G1 informants were more *proficient* than the G2 participants. As a starting point they were considered to be at the same proficiency level (as evidenced by their scores of the proficiency test). In a sense, G2's problem with Norwegian DP internal agreement reflected what Hawkins termed 'selective syntactic fossilization'. These areas were hypothesized to cause persistent problems for L2 learners irrespective of their proficiency levels.

6.3 The experiment

Norwegian is among the few languages that have a 'double definiteness' construction, where there is visible agreement between the prenominal determiner (D) and the suffixed determiner on the noun (N) in terms of gender, number, and definiteness.³⁶ It thus offers a good opportunity to investigate the processing of DP internal agreement. In the present study, we used D as primes, and examine whether the gender, number and definiteness cues on D have effects on the recognition of noun targets that are syntactically related to D. In light of the grammatical agreement processing model I proposed, there is no need to assume a neutral baseline condition. Hence in the present study, I opted for discordant vs. concordant conditions for all the three DP internal agreement features, namely, gender, number and definiteness. If significant differences in reaction times (RT) between concordant and discordant cases are found, it will be an indication that the participants are sensitive to the cues on the prime. In this sense, the syntactic priming effect as revealed by the auditory naming technique can tell us something about production, because the learners must make a judgment on the basis of D-N agreement as they repeat the target noun.

Two peculiarities regarding the features of Norwegian need to be brought to attention. The first is that the 'single definiteness' construction, where the suffixed determiner is left out, is normally ungrammatical in Norwegian, but allowed in the formal style of speech and writing. In addition, it is also the norm in Danish, a language on which the Bokmål Norwegian was developed. I have argued earlier that the surface difference between Norwegian and Danish reflects an underlying syntactic contrast between the two languages, that is, Norwegian D is specified with a [*u*DEF] feature, Danish D is not. Accordingly, the 'single definiteness' construction in the formal style of Norwegian and in Danish reflects a distinction between Null morpheme and the absence of morpheme: the suffixed determiner is null in Norwegian, absent in Danish. The second peculiar property of Norwegian is concerned with the distribution of the

³⁶ Other languages that have the double determination like Norwegian are Swedish and Faroese. Danish, which has a close historical link with Norwegian, does not have 'double definiteness' (see Julien 2003, 2005, for further details).

gender type in Norwegian. The common gender, which is the collapsed form of masculine and feminine genders, is over twice as frequent as the neuter gender. It is especially intriguing to see how these two peculiar properties of Norwegian will affect L1 and L2 processing.

The goals of the present study are: (1) to determine whether native speakers are sensitive to DP internal gender/number/definiteness agreement violations when processing Norwegian; (2) to pinpoint whether the asymmetry of gender frequency has an effect on L1 and L2 processing; (3) to determine whether L2 speakers are sensitive to DP internal gender/number/definiteness agreement violations in L2 grammar; (4) to examine whether there is an L1 transfer effect on L2 processing; (5) to assess whether sensitivity (or insensitivity) to the agreement cues is directly related to how well they produce the correct agreements. That is, whether perception parallels production.

The grammatical agreement processing model proposed above predicts that native speakers will show sensitivity to agreement violations across all the three agreement features when processing Norwegian. However, a cautious note has to be made for definiteness-marking. As have been discussed in the previous chapters and above, the ‘discordant forms’, such as *de gamle planer*, exist as grammatical form in Danish and also in the formal style of speech and writing in Norwegian. So the intriguing question is how native speakers are going to react to these ‘discordant forms’. I would predict that the concordant forms and discordant forms would be responded differently by native speakers of Norwegian. If indeed Norwegian D has a [*u*DEF] feature, along with [*u*GEN] [*u*NUM] features, as I have previously argued, there will be an effect of definiteness, no less than the gender- and the number-marking effects.

If L2 speakers employ the shallow parsing route, they are predicted to be insensitive to the agreement cues on D. If on the other hand, L2 speakers employ the same processing route as native speakers, and if we assume that the uninterpretable features ‘drive’ syntactic computation, L2 speakers are predicted to be only sensitive to the part of the grammar where their L1s share similar agreement properties with L2 Norwegian. Specifically, Romance-speaking informants will be sensitive to gender agreement violations; Romance- and English-speaking informants will be sensitive to number agreement violations; no L2 groups will be sensitive to definiteness agreement violations when processing Norwegian.

6.3.1 Method

a. Participants: Participants in the auditory naming task were the same as in the on-line DP production task. They were composed of sixteen L2 participants and fourteen L1 participants. The L2 participants were further divided into subgroups in two ways. Based on their L1s, they were divided into three groups: Romance (5 informants), English (6 informants), and Chinese (5 informants). Based on their performance in the production task, they were halved into two groups: G1 and G2. G1 were made up of four Romance and three English informants, who were natively like performers in the on-line DP production task; G2 were made up of one Romance informant, three English informants, and five Chinese informants; they were non-natively like performers, who produced correct Norwegian DP internal agreement at around chance level. In order to render L1 and G1 participants comparable, seven native speakers were selected to form a sub-control group. The sub-control group were matched on gender and age with the

G1 (mean ages were 43.2 for G1; 41.6 for the sub-control group). Both the control groups and the L2 groups were naive about the goal of the study.

b. *Materials and design*: Stimuli for the experiment were auditory Norwegian DPs of the ‘double definiteness’ construction, which were made up of a prenominal determiner (D), an adjective, and a noun (N) with a suffixed determiner. As the focus of the experiment was on agreement relations between D and N, the role of the attributive adjective was minimized. Thus the same adjective *gammel* ‘old’ was used in all test items. The form of the adjective was invariant (which was *gamle* in all cases). Depending on the grammatical agreement features being manipulated, the test items were divided into three groups: gender-marking, number-marking, and definiteness-marking, each consisting of 36 test items. In each group the test items were further halved into two conditions: concordant and discordant. By concordant, we mean that the determiner has the appropriate morphology that agrees with the noun target. By discordant, we mean that the determiner has a mismatch of a single feature (gender, number, or definiteness) with the noun target. Table 6.1 shows sample stimuli in gender-, number- and definiteness-marking (See Appendix 7, 8 and 9 for a whole set of materials for gender-, number-, and definiteness-marking respectively).

Table 6.1. Sample stimuli in the auditory naming task (mismatched features in bold form)

	CONCORDANT			DISCORDANT		
GEN	den	gamle	bil-en	den	gaml-e	bord-et
	D-COM.SG.DEF.	old	car-COM.SG.DEF	D-COM.SG.DEF	old	table-NEUT.SG.DEF.
NUM	det	gaml-e	kurs-et	de	gaml-e	hus-et
	D-NEUT.SG.DEF.	old	course- NEUT.SG.DEF.	D-PL.DEF.	old	house- NEUT.SG.DEF.
DEF	de	gaml-e	krig-ene	de	gaml-e	plan-er
	D- PL.DEF.	old	war- PL.DEF.	D- PL.DEF.	old	plan-PL.INDEF.

With respect to gender-marking, I assumed the two-way gender system of Norwegian, neuter gender and common (non-neuter) gender. The number of items was equal for each gender and in each condition. Accordingly, there were 18 items in the gender concordant condition; half of the nouns were of common gender, half of neuter gender. So was the case in the discordant condition.³⁷ As have been previously discussed, the gender information in Norwegian is only visible in singular determiners (*den* for common gender; *det* for neuter); in the plural, gender information on the determiner is neutralized (*de* for both genders). So the nouns and the determiners in the gender-marking group were all in singular form.

In the number-marking group, the noun targets were all in the singular form. The number on the determiners was either singular (in which case the determiners matched the noun targets in number), or in plural (in which case the determiners mismatched the

³⁷ But one test item whose word target is ‘fange’ (which had been planned to be in the discordant condition) was excluded in our final data analysis, as both *fangen* ‘the prisoner’ and *fanget* ‘the lap’ are possible, with differences in meaning.

noun targets in number). For the number concordant condition, we used the same test items as those in the gender concordant condition. Hence the singular D took two forms: *den* or *det*, and there were equal number of cases for *den* and *det*, 9 for each. *De* was used as a prime carrying a plural number cue. There were altogether 18 cases for *de*.³⁸

For definiteness marking group, we opted for plural number for both determiners and noun targets in order to eliminate any interference of gender information in processing (recall that gender information is not visible on D or N in plural). As Norwegian does not have an indefinite plural pronominal determiner comparable to the definite plural pronominal determiner (*de*), I kept the pronominal determiner (*de*) invariant, while using noun endings (*-ene* or *-er*) to signify feature match or mismatch. Accordingly, noun targets with *-ene* suffix were concordant with *de*, because the suffix marks definiteness; noun targets with *-er* suffix were discordant with *de*, because the suffix marks indefiniteness.

All in all, 90 different nouns were selected for this study. There was no repetition of nouns in the test items. Care had been taken to counterbalance the word frequency (based on *the Oslo Corpus of Tagged Norwegian Text*) and the syllable length of nouns in discordant and concordant conditions for all the three feature groups. This was achieved in gender and definiteness marking groups, but in number-marking, only word frequency was counterbalanced, at the cost of syllable length. See Table 6.2 for the mean word frequency and Table 6.3 for mean syllable length of nouns in each cell.

Table 6.2. Mean word frequency of the target nouns in each cell

Determiners	Grammatical agreement features		
	gender	number	definiteness
Concordant	3017	3017	2812
SD	(2408)	(2408)	(1426)
Discordant	3184	2981	2899
SD	(2359)	(1612)	(1443)

Table 6.3. Mean syllable length of the target nouns in each cell

Determiners	Grammatical agreement features		
	gender	number	definiteness
Concordant	2.28	2.28	3.06
SD	(.461)	(.461)	(.938)
Discordant	2.22	2.89	3.22
SD	(.428)	(.758)	(.428)

T-tests for independent samples showed that the mean word frequency differences between concordant and discordant situations in each feature group were not significant [$t_{\text{gen}}(33) = .208$, $p = .837$; $t_{\text{num}}(34) = .185$, $p = .855$; $t_{\text{def}}(34) = .179$, $p = .859$];³⁹ t-tests for independent samples were run to test for a difference between feature groups. No significant differences were found (in all cases, $p > .05$). The mean syllable length

³⁸ The phonetic forms of the three determiners are: /de/ for *det*; /den/ for *den*; /di:/ for *de*.

³⁹ For all the statistic analyses in this paper, alpha was set at .05 level.

differences between concordant and discordant situations were not significant in the gender and the definiteness groups [$t_{\text{gen}}(33) = .375$, $p = .710$; $t_{\text{def}}(34) = .686$, $p = .497$], but significant in the number group [$t_{\text{num}}(34) = 2.922$, $p = .006$], with the average syllable length in discordant cases longer than that in concordant cases. In addition, the average syllable length of the noun targets in definiteness marking was significantly longer than those in gender- and number-marking (in both cases, $p > .05$).

All the test items were audio-recorded by a female native speaker of Norwegian with standard Bokmål pronunciation at a natural speed. Recordings were made in a sound-treated studio and all stimulus preparations were done by using Cool Edit Pro[®] and Praat (Boersma and Weenink, 2006). In preparing the stimuli, one token of each determiner *den*, *det* and *de*, and one token of the adjective *gamle*, and all the noun targets were spliced out. New determiner-adjective pairs (*den gamle*, *det gamle*, *de gamle*, respectively) were formed with the chosen adjective and the determiner exemplars. Each new determiner-adjective pair was then added to a noun target that had been preceded by a corresponding determiner-adjective pair in the recording. By so doing we wish to ensure that the noun targets are preceded by the determiner-adjective pairs of a similar duration. Also great care was taken to ensure natural transitions between the words and to achieve appropriate amplitude relations. Noun durations in the recording were also measured. The average noun durations in each cell are shown in Table 6.4 (the durations of the whole set of noun targets were provided in Appendix 7, 8 and 9).

Table 6.4. Mean word durations in the recording (in milliseconds)

Determiners	Grammatical agreement features		
	gender	number	definiteness
Concordant	624	624	675
SD	(86)	(86)	(53)
Discordant	628	689	718
SD	(88)	(117)	(98)

Though the word durations in the discordant conditions were longer than concordant conditions, especially in the number- and the definiteness-marking groups, as one can observe from Table 6.4 above, t-test results show that there were no significant differences between the concordant and discordant conditions across the three features [$t_{\text{gen}}(33) = .149$, $p = .882$; $t_{\text{num}}(34) = 1.899$, $p = .066$; $t_{\text{def}}(34) = 1.629$, $p = .113$]. Thus the noun durations were comparable for concordant and discordant recordings for all the three feature groups.

The test items of the three groups, namely, gender-, number- and definiteness-marking, were all mixed and randomized, with an interstimulus interval set at 3.5 seconds. The order of presenting the stimuli was the same for all the informants.

c. *Procedure*: Informants were tested individually in a sound-proof phonetic lab. They were informed via written instructions that they were going to hear a series of noun phrases of Bokmål variety, and that they would be asked to repeat the nouns after *gamle* as quickly and as accurately as possible. They also learned that there were both grammatical and ungrammatical expressions in the phrases they were going to hear. It

was emphasized that the noun targets should be repeated in the same form as they appeared in the recording; no correction should be made when repeating the nouns.

The test items were presented to the informants one by one via headphones. The informants' vocal responses were recorded on tape via one of the two channels of a DAT-recorder. The audio signal presented to the informants was recorded simultaneously via the other channel. Prior to the experimental session, all informants were asked to complete a practice session with 3 test items, none of which contained target nouns used in the real trials. Often the practice session was repeated until the experimenter made sure that the informants understood the requirements of the task. Much emphasis was put on speed of reacting, so that the informants were working under time pressure. Norwegian was used throughout the testing session. The test took about 9 minutes, and there was a short break every 3 minutes.

d. *Reliability of measuring reaction times*: In preparing reaction times (RT) measurements, both the audio stimulus signal and the informants' responses were copied onto hard disk and stored as two-track files. Using Cool Edit Pro[®], RTs were measured from the onset of the target word to the onset of the participant's vocal response. All measurements were performed by a research assistant, who was a native speaker of Norwegian specialized in phonetics. In order to check the reliability of the data, eleven randomly chosen items from each of two informants were measured independently by another phonetician. The difference between the mean RTs measured by the two phoneticians was only 3.5 milliseconds, and was mainly caused by the use of different measurement criteria for one of the test words. Excluding this case, the mean difference was reduced to 1.1 milliseconds. These results thus indicate that the RTs measured by the research assistant are reliable.

6.3.2 Results

a. *Data trimming*: There were no cases where the L1 or the L2 participants failed to respond. But both groups made a small number of errors. Errors, including false starts, hesitation, failure to produce the target correctly, were excluded from the final data analysis. For the L1 group, error rate for target nouns preceded by determiners *den* and *det* was 3.1% (15 out of a total of 490 trials); for nouns preceded by *de* was 9.5% (72 out of a total of 756 trials). The relatively large error rate for the latter case was mainly due to the fact that some informants automatically corrected the nouns that were discordant with the determiner *de*. So instead of repeating the target noun, say *systemet* 'the system', in the stimulus presented to them, they produced *systemene* 'the systems'. This accounts for 57% (41 out of 72) of the total errors in the *de* cases, which somewhat suggests that native speakers of Norwegian expect the identity of the following noun to be in the form of *-ene* as soon as they hear the determiner *de*.

For the L2 group, error rate for target nouns preceded by determiners *den* and *det* was 3.9% (22 out of a total of 560 trials). One informant, R5, made errors for almost all the items in definiteness-marking group. So her data in this group was discarded. At last, error rate for nouns preceded by *de* was 4.8% (39 out of a total of 810 trials).

In addition, RTs longer than 1200ms were deemed outliers and were removed. These data points were extremely small for both groups: 0.2% (3 out of a total of 1240 trials) for the L1 group, 0.8% (11 out of a total of 1370 trials) for the L2 group. The remaining data points for the fourteen native speakers and for the sixteen L2

participants were fed into SPSS (version 14). The participants' mean RT to each test item was submitted to analysis of variance (ANOVA).

b. *The L1 participants*: A 2(concordant vs. discordant Ds) by 3 (gender, number, and definiteness) ANOVA was conducted. Table 6.5 shows the mean RT in each cell and naming time differences between concordant and discordant cases in each agreement feature group.

Table 6.5. Control group's Mean RTs in each cell (RT in ms)

Determiners	Grammatical agreement features			total
	gender	number	definiteness	
Concordant	503	503	560	522
SD	(51)	(51)	(47)	(56)
Discordant	568	610	632	604
SD	(78)	(70)	(95)	(84)
Difference	65	107	72	82

The ANOVA yielded a main effect of concordance [$F(1, 101) = 38.31, p < .001$]. The strong effect of concordance is in the predicted direction. Collapsed over all three agreement features, targets primed by a concordant D were responded to 82ms faster than those primed by a discordant D, indicating that native speakers are overall sensitive to grammatical agreement violations. The main effect of agreement features was also significant [$F(2, 101) = 7.34, p = .001$]. A Turkey HSD post-hoc test revealed that the average RT to definiteness-marking (596ms) was significantly longer than the average RTs to gender-marking (536ms) and to number marking (557ms); The average RTs to gender-marking and number-marking did not differ, though. As syllable length of noun targets is likely to influence reaction time (cf. Bates, *et al.* 1996), the significantly longer RTs to definiteness-marking can be attributed to its longer average syllable length. As I noted before, the mean syllable length in definiteness-marking was the longest among the three groups.

Separate analyses were carried out for the three agreement feature groups in order to assess the participants' sensitivity to gender-, number-, and definiteness marking, respectively. The 65ms concordance-discordance difference in gender-marking, 107ms difference in number-marking, and 72ms difference in definiteness-marking are all significant [$t_{\text{gen}}(33) = 2.91, p = .006$; $t_{\text{num}}(34) = 5.21, p < .001$; $t_{\text{def}}(34) = 2.86, p = .007$]. Hence clear priming effects were obtained for all the three agreement features. These results show that for gender, number and definiteness marking, the cue type (concordant vs. discordant) on D influences native speakers' processing of a subsequent noun: a concordant D speeds up auditory naming times as compared against a discordant D. These results clearly indicate that L1 speakers are sensitive to gender, number and definiteness cues when processing their native language. The predictions about L1 processing based on the agreement process model were thus borne out.

Especially notable was the robust concordance-discordance difference in definiteness-marking. As previously discussed, native speakers of Norwegian actually use in their formal style of speech and writing such a construction as *de gamle planer*

‘the old plans’, which was deemed ‘discordant’ form in this experiment. Despite their exposure to the alternative form, an effect of definiteness was observed with a size similar to those of gender- and number-marking effects. This strongly suggests that native speakers of Norwegian distinguish concordant form from the discordant one and that the ‘double definiteness’ construction was taken as the norm and the ‘single definiteness’ was considered as an anomaly in Norwegian.

Turning to number-marking, it should be recalled that both word frequency and syllable length were counterbalanced for discordant and concordant conditions in gender and definiteness marking. But for number marking, only the word frequency was counterbalanced, whereas there was a significant difference between concordant and discordant cases with respect to syllable length. In order to eliminate the possible influence of syllable length on the t-test result of number-marking, I selected 11 items each from the discordant cases and concordant cases, this time having both word frequency [$M_1=2178$, $M_2=2224$; $t(20)=.09$, $p=.929$] and syllable length (mean syllable length= 2.45 in both cases) counterbalanced. I then ran a (two-tailed) t-test for independent samples with the new data, and again the number-marking effect was found [$t(20) = 3.06$, $p=.006$]. It was thus confirmed that the number-marking effect was obtained, independent of word frequency or syllable length.

Now we look more closely on how native speakers react to gender marking when processing Norwegian. As I noted before, there is an asymmetry of distribution in the gender system of Norwegian, with common gender (*den*) twice as frequent as neuter gender (*det*). We wish to ascertain whether this asymmetry of gender type distribution has an effect in L1 processing. If we place the influence of agreement processing on a post-lexical, syntactic mechanism, we expect that common gender and neuter gender be treated equally, the asymmetry of gender distribution having little effect on processing. If, on the other hand, this information is utilized at a lexical level, we expect that neuter gender nouns will be easier to process, because the neuter gender activates a more restrictive noun set, which may facilitate lexical search.

A 2(concordant vs. discordant) by 2 (common vs. neuter) ANOVA was performed to test for a difference in RTs between common and neuter gender in both concordant and discordant conditions. As expected, the ANOVA revealed a clear gender priming effect [$F(1, 31) = 7.97$, $p=.008$]; separate analyses showed significant priming for both common gender nouns [$t(241) = 2.714$, $p=.007$], and neuter gender nouns [$t(211) = 2.66$, $p=.008$]. However, there was no effect of grammatical gender (common vs. neuter) [$F(1, 31) = .061$, $p=.806$], and no interaction between cue type (concordant vs. discordant) and grammatical gender (common vs. neuter) [$F(1, 31) = .057$, $p=.812$]. As can be observed from Figure 6.1, the green line, representing the neuter gender nouns and the blue line, representing the common gender nouns, nearly overlap. The mean RTs to common and neuter gender nouns were almost the same (both 568ms) in the discordant condition and very close (only approximately 11ms difference) in the concordant condition. The results indicate that neuter and common genders were treated equally by native speakers, despite the asymmetry across the gender type. This is expected if we place the influence of agreement processing on a post-lexical, syntactic mechanism. We can then attribute the gender priming effect mainly to a syntactic mechanism. The lexical activation, if any, is not obvious in L1 processing in the present study.

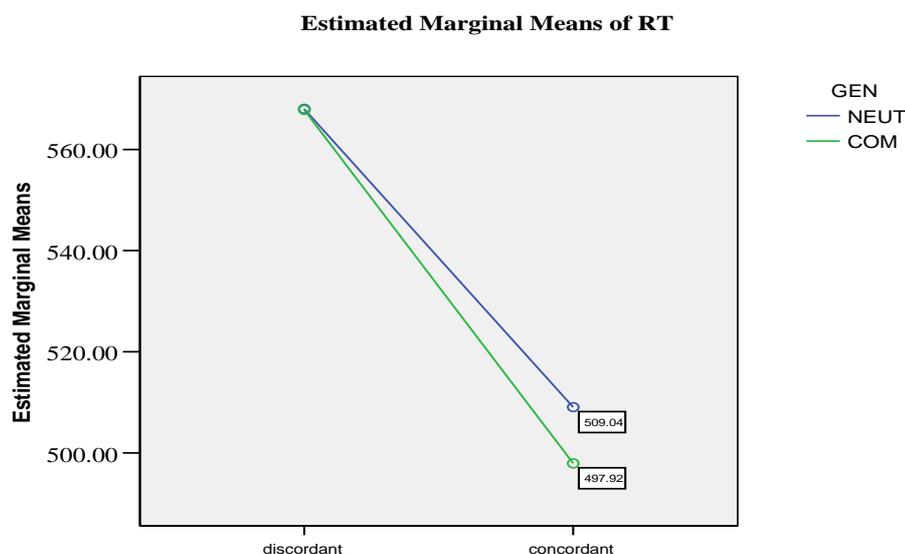


Figure 6.1. The L1 group's mean RTs to concordant and discordant neuter and common gender-marking

Now let us turn to the sub-control group and examine whether the same priming effects will be retained when we lessen the number of participants. The same data analysis procedure was applied to the sub-control group, which was made up of seven L1 participants. In calculating the RTs in number marking, only the 11 data points in each condition (with syllable length and word frequency counterbalanced) were used. By so doing, I wish to eliminate the possible interference of both syllable length and word frequency. The sub-control group's the mean RT in each cell and the average naming time difference in each agreement feature group are shown in Table 6.6.

Table 6.6. The sub-control group's mean RTs in each cell (RT in ms)

Determiners	Grammatical agreement features			
	gender	number	definiteness	total
Concordant	569	582	611	588
SD	(53)	(61)	(48)	(55)
Discordant	647	644	682	659
SD	(80)	(70)	(74)	(76)
Difference	78	62	71	71

The ANOVA again yielded a main effect of concordance, [$F(1, 87) = 25.47$, $p < .001$]. Collapsed over all the three agreement features, targets primed by a concordant D were responded to about 71ms faster than those primed by a discordant D. The (two-tailed) t-tests for independent samples show that the 78ms concordance-discordance difference in gender-marking, 62ms difference in number-marking, and 71ms difference in definiteness-marking are all significant [$t_{\text{gen}}(33) = 3.41$, $p = .002$; $t_{\text{num}}(20) = 2.22$, $p = .038$; $t_{\text{def}}(34) = 2.78$, $p = .009$]. Hence the sub-control group reacted to agreement violations in

the same way as the control group. This demonstrates that the priming effects in L1 processing are robust, not being weakened by the smaller number of participants in the sub-control group.

In summary, it has been found that the native speakers were sensitive to the gender, number, and definiteness agreement cues on D when processing Norwegian, as evidenced by the significant concordance-discordance differences in each agreement feature group. For gender-, number-, or definiteness-marking, RTs to nouns preceded by concordant Ds were significantly faster relative to nouns preceded by discordant Ds. The main findings for L1 processing remained the same regardless of the sample size (n=14 participants in the control group; n=7 participants in the sub-control group). These findings are consistent with the predictions based on the agreement processing model I have proposed.

c. The L2 participants: The data analysis procedure for the L2 participants is the same as I used for L1 participants. In calculating the RTs in number marking, again I only used the 11 data points, with both syllable length and word frequency counterbalanced, in order to rule out the possible interference of syllable length and word frequency.

We are firstly concerned with the question whether or not the L2 participants as a whole are sensitive to grammatical agreement violations when processing L2 Norwegian. To explore this issue, the L2 participants' mean RT to each test item was submitted to analysis of variance (ANOVA). Table 6.7 shows the mean RT in each cell and the average naming time difference in each agreement feature groups.

Table 6.7. The L2 informants' mean RTs in each cell (RT in ms)

Determiners	Grammatical agreement features			total
	gender	number	definiteness	
Concordant	602	602	638	616
SD	(54)	(67)	(45)	(56)
Discordant	611	619	671	638
SD	(56)	(38)	(51)	(57)
Difference	9	17	33	20

The ANOVA did not yield a main effect of concordance [$F(2, 87) = 3.12, p = .081$]. Collapsed over all the three agreement features, targets primed by a concordant D were responded to only about 20ms faster than those primed by a discordant D. This indicates that the L2 participants were overall insensitive to grammatical violations in the L2 grammar. The (two-tailed) t-tests for independent samples showed that the 9ms concordance-discordance difference in gender-marking, and 17ms difference in number-marking was not significant [$t_{\text{gen}}(33) = .479, p = .635$; $t_{\text{num}}(20) = .710, p = .486$]. But the 33ms difference in definiteness-marking was (weakly) significant [$t_{\text{def}}(33) = 2.08, p = .044$]. In appearance, these results seem to indicate that the endstate L2 speakers of Norwegian were insensitive to gender- and number-marking, but were sensitive to definiteness-marking. This concordance-discordance difference in definiteness marking obviously runs counter to my predictions. Since none of the source languages have the [$u\text{DEF}$] feature, our agreement processing model would predict that

it would be least likely for the L2 speakers to be sensitive to definiteness marking when processing L2 Norwegian. The possible reason for this apparent definiteness effect in L2 processing will be analyzed in the discussion section.

As a parallel to data analysis in L1 processing, I also examine the effect the asymmetry of gender type distribution in L2 processing. We have known from the preceding chapter that L2 speakers tend to use the common gender as a default. We are thus interested in exploring whether the common gender nouns are earlier to process relative to the neuter gender nouns for these L2 participants.

A 2(concordant vs. discordant) by 2 (common vs. neuter) ANOVA was conducted. As expected, the ANOVA did not reveal any gender priming effect [$F(1, 31) = .239$, $p = .628$], indicating that L2 participants were overall insensitive to gender agreement violations in the L2 grammar. There was no effect of grammatical gender (common vs. neuter) [$F(1, 31) = 1.001$, $p = .325$], and no interaction between cue type (concordant vs. discordant) and grammatical gender [$F(1, 31) = .182$, $p = .672$]. As we can observe from Figure 6.2, the average RTs to neuter gender nouns were extremely close in the two conditions (616ms in the discordant condition vs. 615ms in the concordant condition); the difference between the average RTs to common gender nouns in the two conditions was also small (approximately 14ms). We can also observe that in both conditions, the blue line, representing RTs to neuter gender nouns, lies above the green line, representing RTs to common gender nouns. This means that the average RTs to neuter gender were longer than the RTs to common gender in both concordant and discordant conditions, indicating that neuter gender nouns were generally more difficult for the L2 participants to process than common gender nouns. However, T-tests for independent samples showed that the RT differences between the two cells did not reach a significant level (in all cases, $p > .05$).

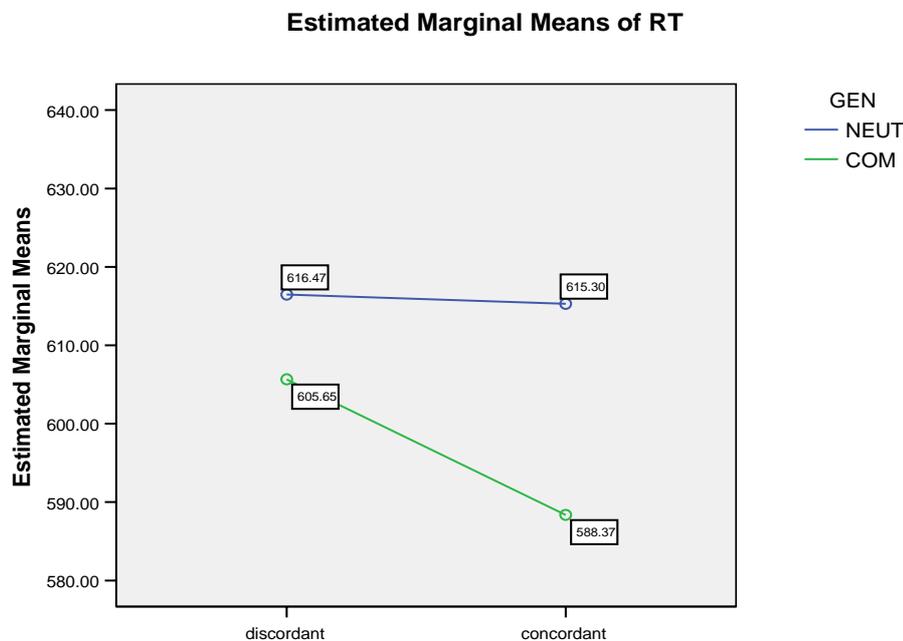


Figure 6.2. The L2 participants' mean RTs to concordant and discordant neuter and common gender-marking

In order to investigate the L1 transfer effects on L2 processing, we need to separate the L2 participants based on the fact of whether their L1s share the similar agreement properties with L2 Norwegian. For gender agreement, only the Romance languages are syntactically similar with Norwegian; for number agreement, both English and the Romance languages are syntactically similar with Norwegian. So we look at the L1 transfer effect on gender-marking using the data of the Romance group, on number-marking using the data from both the Romance group and the English group. If L1 transfer effect occurs in L2 processing, our agreement processing model will predict that the Romance group will be sensitive to gender agreement violations, and that Romance group and the English group will be sensitive to number agreement violations. However, the t-tests for independent samples showed that concordance-discordance difference in neither agreement feature group was significant [$t_{\text{gen}}(164) = .465$, $p = .642$; $t_{\text{num}}(230) = .669$, $p = .504$].⁴⁰ So contrary to our predictions, the L2 participants showed no nativelike sensitivity to agreement violations in L2 grammar, suggesting that L1 transfer did not occur in L2 processing.

The last issue we would like to investigate is the correlation between production and perception. In particular, I ask the question whether sensitivity (or insensitivity) to the agreement cues is directly related to how well they produce the correct agreements. In order to address this issue, I look at the data from G1 participants (nativelike performers in the production task) and G2 participants (nativelike performers in the production task) separately. Table 6.8 and 6.9 give the mean RTs in each cell for G1 and G2, respectively.

Table 6.8. G1 L2 participants' mean RTs in each cell (RT in ms)

Determiners	Grammatical agreement features			total
	gender	number	definiteness	
Concordant	565	569	606	582
SD	(58)	(66)	(55)	(61)
Discordant	572	611	629	603
SD	(75)	(60)	(74)	(73)
Difference	7	42	23	21

Table 6.9. G2 L2 participants' mean RTs in each cell (RT in ms)

Determiners	Grammatical agreement features			total
	gender	number	definiteness	
Concordant	576	629	598	597
SD	(67)	(74)	(59)	(67)
Discordant	583	642	632	616
SD	(66)	(67)	(75)	(74)
Difference	7	13	34	26

⁴⁰ Due to relatively small number of participants, all the data points (instead of the mean RT to each test item) were used in running the t-tests.

Two striking facts can be observed from the data above. First, the overall priming effects were quite comparable between the two L2 groups. Collapsed over all the three agreement features, the concordance-discordance difference was 21ms for G1, and 26ms for G2. It seems that nativelike production of Norwegian DP internal agreement is not directly related to how well L2 learners use agreement cues in perception. Second, for both L2 groups, the mean naming time differences between the concordant and discordant conditions were much smaller than those of the native controls, suggesting that L2 participants can be much less sensitive to the agreement cues on D. The ANOVA confirmed these observations. There was not a main effect of concordance in either G1 ($F(1, 86) = 2.947, p = .090$) or G2 ($F(1, 87) = 2.602, p = .110$). Separate analyses for the three agreement feature groups did not show any significant concordance-discordance difference in any feature groups for either of the two L2 groups [in all cases, $p > .05$]. These results thus clearly indicate that neither G1 nor G2 participants were sensitive to the agreement cues on D, whether or not they performed nativelike in the production task.

6.4 Summary of findings and discussion

6.4.1 A summary of L1 & L2 processing results

The present experiment used an auditory naming technique to examine how L1 and adult L2 speakers of Norwegian reacted to violations of DP internal gender/number/definiteness agreement when processing Norwegian. The main findings emerged from the experiment are summarized below:

- The native speakers of Norwegian were sensitive to gender, number, and definiteness agreement cues on D, as evidenced by the significant concordance-discordance differences in RTs to gender-, number, and definiteness-marking; these priming effects were equally robust for the control group and the sub-control group, irrespective of a difference in sample sizes of the two groups.
- Particularly notable is the effect of definiteness, which was no smaller in size than gender- and number-marking effects, despite the fact that the discordant form exists in Danish and in the formal style in Norwegian.
- In L1 processing, the concordant effect in gender-marking does not interact with gender type; the RTs to common gender nouns and to neuter gender nouns were found to be extremely close, even though the genders differ dramatically in their frequency.
- The L2 speakers were overall insensitive to DP internal agreement violations in L2 grammar; only in the definiteness-marking group was there a significant concordance-discordance RT difference, but this significant difference was not observed in either G1 or G2.
- The asymmetry of distribution in the gender system of Norwegian (with common gender twice as frequent as neuter gender) had little effect in L2 processing, as the difference in RTs to common gender and to neuter gender were not significant; however, RTs to neuter gender nouns were slower than

RTs to common gender nouns, indicating that for the L2 participants, neuter gender nouns were a little more difficult to process.

- Although the Romance languages share with Norwegian similar gender agreement properties, and Romance and the English languages share with Norwegian similar number agreement properties, the Romance group were not sensitive to gender-marking, nor were the Romance and the English groups combined sensitive to number-marking, suggesting that a different processing mechanism might be involved in L2 processing.
- The nativelike performers did not behave differently from the non-nativelike performers; both showed insensitivity to all kinds of grammatical agreement violations in the L2 grammar. This is a clear indication that the L2 participants' insensitivity to the agreement cues is not directly related to how well they produce the correct agreements, in other words, perception does not parallel production.

These findings raise some interesting questions in need of explanation. First of all, where is the locus of priming effects: a lexical locus or a post lexical syntactic feature checking? Second, how does the gender/number/definiteness information influence the word recognition process in L1 processing? Thirdly, is there a L1 transfer effect in L2 processing? And lastly, why are L2 speakers insensitive to grammatical violations in L2 grammar? These issues will be addressed based on the observations from the experiment, and where it is necessary, I will draw on empirical results from relevant studies in literature.

6.4.2 The locus of priming effects in L1 processing

Generally, the findings from L1 processing support the argument that the priming effect is due to syntactic feature checking, not lexical access. Two aspects of the findings emerged from the experiment support this view. First, it was observed that the common gender and the neuter gender on the prime provided equivalent facilitation effects, even though the common gender is twice as frequent as the neuter gender. Thus the asymmetry in the distribution of gender type did not provide an additional advantage for the more restrictive neuter gender. This finding suggests that the influence of gender-marking plays little, if any, role in constraining lexical activation space. Second, not only gender-marking effect, but also number- and definiteness-marking effects were observed. Number and definiteness, unlike gender, are not properties inherent in the noun, so lexical activation cannot accommodate number- and definiteness-marking effects. We thus have to conclude that it is a post-lexical, syntactic mechanism that contributes the observed effects. To the best of my knowledge, no previous studies have investigated L1 processing of grammatical agreement violations involving gender, number, and definiteness features. The uniform performance pattern across these grammatical features exhibited by the native speakers thus provides a strong piece of empirical evidence for the argument that the effect is due to a syntactic mechanism.

If we consider the on-line nature of the task, which requires the informants to respond as quickly as possible, we wonder why the native speakers simply could not ignore the agreement cues on the primes altogether. They could surely observe that the primes offered completely unreliable information about the identity of target nouns. It should be recalled that in our experiment all the materials were mixed and randomized

before being presented to the informants. Hence the primes on D took three forms: *det* (neut. sg.), *den* (com. sg.) and *de* (pl.), and depending on which groups they appear, the target nouns could carry a variety of suffixes: *-en*, *-et*, *-ene* or *-er*. Given these complications, it is not predictable which prime will be in the next test item or which suffix the noun target will take. Then why, as it turned out, is the information on D “hard to suppress, though it would be convenient to do so” (Bates *et al.* 1996)? If we assume that syntactic feature checking takes place in L1 processing, these findings are naturally interpretable. Details of feature processing mechanism will be provided in the next section.

Although the results obtained from L1 processing overwhelmingly point to the involvement of a syntactic mechanism in grammatical agreement processing, I do not exclude the possibility that gender information may influence lexical activation. Since parsing is incremental and gender information (in the gender-marking group) has been provided by a prenominal determiner which will be heard prior to the noun targets, it is highly likely that lexical activation occurs as soon as a gender-marked D element (*den* or *det*) is discovered. Consequently, a noun set with the same gender will be immediately activated. A more restrictive noun set is supposed to facilitate lexical research compared with a larger noun set. But what might account for the ineffectiveness of gender distribution in affecting lexical activation in L1 processing of Norwegian? A possible reason could be that the gender information in Norwegian is not only available in the determiner, but also explicitly marked on the noun suffix, as in *det gamle huset*. Thus attention will naturally be directed to the end of the noun, which weakens the initial facilitation effect due to lexical constraint. In this sense, lexical activation has given way to the syntactic mechanism, which implies an explicit ‘checking’ of word endings.

6.4.3 Agreement processing mechanisms

The observation that native speakers of Norwegian showed sensitivity to DP internal gender/number/definiteness agreement violations when processing Norwegian is fully in line with the predictions based on the agreement processing model. The model was proposed with an attempt to answer where and how the gender/number/definiteness cue on D influences the word recognition process. In particular, this processing model incorporates the grammatical agreement mechanism developed in the Minimalist Program. The idea is that agreement involves a Probe-Goal relation between a head that carries uninterpretable formal features and a constituent that has the matching interpretable features. Once the Probe-Goal relation has been established, the feature values on the Goal are copied onto the Probe. The uninterpretable features on the Probe are thus valued and deleted. The newly gained feature values will then be sent to phonology and spelt out as appropriate morphology. This process has been illustrated using the Norwegian example *det gamle huset* ‘the old house’, here repeated in (6.5).

It is assumed that the uninterpretable features drive the parsing process so that automatic syntactic computation takes place in processing as it does in formulating an utterance. In L1 processing, I argued that an elementary tree of DP will be selected as soon as a D element is discovered. Once a DP is analyzed as such, the Agree operation is forced to apply if D carries an uninterpretable feature. Feature valuation will take place between the probe D and the goal N. Processing will be efficient if all features are matched; in case a single mismatch of features is detected, attention will be directed to

intense dialogue between psycholinguists and syntacticians. I fully agree with the authors on this point, and believe that more research should be carried out to bridge the current gap between linguistic theories and data from psycholinguistic experiments.

6.4.4 L1/L2 processing differences

One of the important findings from this experiment is that endstate L2 speakers of Norwegian reacted to L2 DP internal grammatical agreement violations distinctly from native speakers. While native speakers showed sensitivity to gender/number/definiteness cues on D, L2 speakers were overall insensitive to grammatical violations in the L2 grammar. The L2 speakers' performance in this task was not influenced by the properties of their L1s, nor their implicit grammatical knowledge about Norwegian DP internal agreement. It was observed that Romance group were not sensitive to gender agreement violations in the L2 grammar, and the Romance and the English groups combined did not show sensitivity to number agreement violations either. Moreover, G1 and G2 informants displayed a similar processing pattern, though the two groups differed as to whether they had nativelike performance with respect to Norwegian DP internal agreement.

If sensitivity to grammatical agreement violations is a consequence of automatic feature checking in a c-command configuration as has been argued above, the L2 speakers' overall insensitivity to agreement violations in the L2 has to be an indication that L2 processing is not automatic in nature. This means that L2 speakers process Norwegian DPs without invoking full syntactic computation. In this case, it is the Probe-Goal relation between D and N that is not established; hence the information on D presumably has no effects on the processing of N. This finding thus lends support to the *shallow structure hypothesis* (SSH, Clahsen and Felser 2006a, 2006b), according to which L2 learners underuse structural information when processing the target language. If the L2 speakers used the shallow processing route, they would be able to ignore the information on D altogether, focusing on the lexical-semantic and phonetic information of target nouns instead.

In the following I will show that assuming that L2 speakers are more influenced by the semantic and phonetic features will be helpful in accounting for the observed subtle, but nonetheless interesting, differences between L1 and L2 processing. First, while L1 speakers process common gender nouns and neuter gender nouns with equal efficiency, L2 speakers were shown to process common gender nouns faster than neuter gender nouns. Note that L2 speakers' difficulty processing the neuter gender words was not related to the gender cue type (concordant vs. discordant) on the preceding D. The average RTs to neuter nouns were almost the same in the two conditions, and in both conditions, they were longer than RTs to common gender nouns. As the common gender nouns in the experiment carried an *-en* suffix, the neuter gender nouns carried an *-et* suffix, this observation can be reinterpreted to the effect that the nouns with an *-et* suffix were more difficult to process than those with an *-en* suffix. This result from L2 processing, I suggest, is likely to be related with the observation from the gender-assignment task that L2 speakers of Norwegian usually use common gender as a default. Thanks to the default gender strategy L2 speakers utilize in their L2 production, definite nouns marked with a common gender *-en* are supposedly more frequent than nouns marked with a neuter gender *-et* suffix. Though the noun frequency in the experiment was counterbalanced, the nouns with *-en* suffix enjoy higher frequency than the nouns

with *-et* suffix in the interlanguage grammar of the L2 speakers. Thus it is much easier for the L2 speakers to identify nouns with a *-en* suffix as a word, whereas nouns with a *-et* suffix probably require a more detailed analysis before they were identified as a word. As a result, the L2 speakers needed some extra time to repeat nouns with a neuter suffix.

Second, the effect of definiteness in L1 processing was robust and constant, not being dramatically influenced by the sample size (that is, a clear effect of definiteness was retained in the sub-control group). By contrast, in L2 processing only a weak definiteness-marking effect was observed when all the L2 data points were included. When we looked at G1 and G2 separately, however, the priming effect disappeared, though RTs to noun targets with *-er* suffix were still longer to RTs to noun targets with *-ene* suffix in both groups. The observation that L2 speakers showed sensitivity to definiteness marking obviously runs counter to my predictions. As none of the source languages have the [uDEF] feature, the agreement processing model would predict that it would be least possible for the L2 speakers to be sensitive to definiteness marking. Then what might explain the L2 participants' apparent sensitivity to definiteness agreement violations in the L2 grammar?

I suggest that the reason for the significant concordance-discordance difference in definiteness-marking may lie in a phonetic level, rather than a syntactic level. The Norwegian indefinite plural suffix *-er* is pronounced as a schwa plus /r/. The Norwegian /r/, which is an apical trill, is notoriously difficult for the L2 learners to acquire, because it involves phonetic features that are unused in the L2 participants' mother languages.⁴¹ These phonetic features are probably unavailable to L2 learners beyond a critical period (cf. Brown 2000). According to my research assistant, who was a native speaker of Norwegian with a strong background in phonetics, the way the L2 participants pronounced the Norwegian /r/ was generally not nativelike. By this sound, it is easy to distinguish 'foreign' speakers from native speakers. The L2 participants themselves were aware of this according to the pre-experimental interviews with them (cf. Section 5.2.1). Hence it is very likely that the L2 participants' difficulty in pronouncing the Norwegian /r/ result in slightly longer repeating time for the target nouns with an *-er* suffix, which in turn render some difference between the RTs to noun targets with *-ene* suffix and *-er* suffix. But since the auditory naming task requires that the informants should respond as quickly as they can, there could be a trade-off between the lengthened repetition time for nouns with an *-er* suffix and the highly speeded responding. As a result of this, the difference between the RTs to noun targets with *-ene* suffix and *-er* suffix could not be robustly significant. As we see, no significant difference was observed when G1 and G2 were examined separately, thus lending further support for my argument that the observed concordance-discordance difference was due to a phonetic reason (specifically, the L2 participants difficulty in pronouncing the Norwegian /r/ sound in the *-er* suffix), rather than a syntactic one.

Note however, the SSH differentiates between the domain of morphology and the domain of syntax. For morphological processing, L2 learners are claimed to be able to employ the same mechanism as native speakers; for sentence processing, however, L2 learners are claimed to rely more on lexical-semantic ("shallow") cues rather than

⁴¹ Using phonetic features one can specify the English, Chinese, Italian/Spanish /r/ as an alveolar approximant, the Norwegian one in such a postvocalic position as an apical trill, according to van Dommelen (personal communication, November 2006),

syntactic information. Morphosyntactic agreement is said to belong to the domain of morphology, therefore nativelike processing is predicted to be possible (cf. Clahsen and Felser 2006b:111). According to this argument, we would expect the L2 speakers in the present study to be sensitive to DP internal agreement violations in the L2 grammar, contrary to what we found. The present finding that the L2 speakers behaved differently from the native speakers suggests that they had used a different processing route. Thus, the view that morphosyntactic agreement belongs to the domain of morphology is obviously misguided. As I have argued before, the internal structure for morphosyntactic agreement is not as shallow as inflected words. Under current minimalist assumptions, both morphosyntactic agreement and “non-local” (e.g. subject-verb) agreement involve the checking of formal features between two constituents that are in a hierarchical syntactic structure (i.e. in a c-command configuration). It is not obvious why formal feature checking is possible for morphosyntactic agreement, whereas impossible for non-local agreement. Moreover, if ‘local’ means two elements being adjacent to each other, morphosyntactic agreement between D and N is not ‘local’ at all due to the existence of several intervening heads in accordance with Julien’s (2003, 2005) DP model. In short, there are both empirical evidence and theoretical justifications for putting morphosyntactic agreement (at least the D-N agreement under investigation) into the domain of syntax. The L1/L2 performance differences suggest that L2 speakers employed different processing routes: shallow processing route for L2 speakers, full parsing route for native speakers.

Clahsen and Felser (2006b) offer an explanation why L2 learners are restricted to shallow processing in the domain of syntax. Assuming that parsing mechanisms are universal and are available to L2 processing, they attribute L2/L1 processing differences to inadequacies of the L2 grammar, and predict that L2 learners can develop nativelike parsing abilities after acquiring a nativelike grammar. The results from the experiment, however, provide evidence against this prediction. Before the experiment investigating L2 processing, the L2 speakers’ grammatical knowledge regarding Norwegian DP internal agreement had been tested using an on-line task. Particularly we found that the G1 participants had nativelike performance. If on-line tasks are able to “tap” the learners’ implicit linguistic knowledge, as has been generally believed, the finding would imply that the G1 participants had developed nativelike linguistic competence with respect to Norwegian DP internal agreement. That is to say, they had specified [*u*NUM], [*u*GEN], and [*u*DEF] features on D in their interlanguage grammar. However, the L2 processing results suggested that syntactic computation was not triggered, despite the similarity in the grammatical representations between G1 informants and native speakers. This observation suggests that, contrary to the view that L1/L2 processing difference is due to inadequacy of L2 grammar, L2 speakers may have not completely acquire the same processing mechanism as the natives, even after they had developed nativelike grammatical knowledge. As such, the present study supports the view that the age of acquisition is a crucial factor in L2 processing; the availability of the full parsing route to L2 learners may be subject to a critical period (see for example, Weber-Fox and Neville 1996; Hahen, and Friederici 2001; Ullman 2004).

The observed discrepancy between the learners’ nativelike knowledge of Norwegian DP structures and their processing performance indicate that perception does not parallel production – L2 speakers cannot use agreement cues on D in perception, even if they are able to produce Norwegian DP internal agreement in a nativelike fashion. I can

conclude then that L2 speakers' insensitivity to the agreement cues is not directly related to how well they produce correct DP-internal agreements.

6.4.5 Is there an L1 transfer effects on L2 processing?

The role of L1 transfer in L2 processing is an issue of great controversy. Clahsen and Felser (2006) claim that L1 transfer influences L2 processing only indirectly, as a consequence of one or more of the knowledge sources that feed the processing system being affected by properties of the L1. This claim has been attested by several studies on parsing of complex grammatical structures in the L2 (e.g. Sabourin 2003; Marinis, Robert, Felser and Clahsen 2005), but the available literature presents a mixed picture for the role of L1 transfer in L2 processing of local domain mismatches. For example, in an ERP study, Sabourin (2003) examined how proficient German-, Romance- and English-speaking learners reacted to gender agreement violations in L2 Dutch. The P600 response was observed only in the German group, but not in the Romance or the English group. This result has been interpreted by the author as evidence showing L1 transfer in L2 processing, because she argues that among the source languages only German has a gender system that is congruent to Dutch gender system (that is, the nouns in the two languages share the same gender). At this juncture, a methodological question arises as to what count as 'similarity' between L1 and L2. In case of Sabourin's study, does the similarity mean the two languages share the same syntactic features, or they have to be additionally congruent in the gender system? As has been shown, Romance languages also display D-N gender agreement. If we take the former criterion of similarity between L1 and L2, Romance languages and German are alike in involving a [*u*GEN] feature on D. So if L1 transfer is involved in L2 processing as Sabourin argues, we should expect Romance-speaking learners to behave like Dutch native speakers as well, contrary to what she found. Clahsen and Felser pointed out that the German informants' nativelike processing could be attributed to the higher proficiency of the German group rather than L1 influence, as the German-speaking learners were also the only ones who had demonstrated above-chance sensitivity to Dutch gender concord in a judgment task. They are implying that L1 transfer influences L2 processing only in highly proficient L2 speakers. However, whether L2 proficiency influences processing is also an open question. ERP studies reveal conflicting evidence for the role of L2 proficiency in processing. For example, Gillon-Dowens, *et al.* (2004) showed that competent English-Spanish bilinguals are sensitive to number violations, but not to gender violations in sentence contexts. The author concluded that whether or not nativelike L2 processing can be obtained will depend on proficiency and similarity between L1 and L2. Tokowicz and MacWhinney (2005), on the other hand, made an opposing observation. They found that very low-proficient English-speaking learners of L2 Spanish were implicitly sensitive to gender agreement violations but not to number violations in a grammatical judgment test. Their finding seems to suggest that neither proficiency nor similarity between L1 and L2 are relevant in determining nativelike L2 processing. In any case, final conclusions regarding the role of L1 transfer and L2 proficiency in L2 processing remain yet to be made. These studies together seem to show that task differences and coherence of languages are also factors influencing L2 processing.

The results from the present study point to no L1 transfer on L2 processing. As have been shown, the Romance and the English group combined did not show sensitivity to

number violations and the Romance group did not show sensitivity to gender violations, despite the similarity in the respective formal features between L1s and the L2. The L2 speakers' insensitivity to agreement violations in the L2 grammar is not related to the availability of the uninterpretable features in the learners' L2 grammar. The L2 speakers' performance in the on-line production task suggested that the uninterpretable features present in the L2 speakers' respective L1s were also available in their interlanguage grammars. But crucially these uninterpretable features did not trigger feature checking, indicating that L2 processing may not involve full syntactic computation. Rather, L2 processing is guided by lexical-semantic information, as the SSH claims. Considering that L1 transfer influences L2 processing depending on task differences, a cautious note to be taken is that so far I am not certain whether this phenomenon is specific to the auditory naming task or not. Future research involving many other languages and across a variety of tasks will be conducted in order to reach more reliable conclusions.

6.5 Conclusion

The present study used an auditory naming technique to investigate how gender/number/definiteness information on D influences the processing of the subsequent noun in Norwegian as first and second language. The results revealed contrasting performance between native speakers and the L2 speakers. Effects of gender-, number- and definiteness-marking were observed in L1 processing, but not in L2 processing, which indicates that the native speakers were sensitive to the agreement cues on D, whereas the L2 speakers were not. The results obtained from L1 processing provided evidence for the syntactic nature of the priming effects, and can be accommodated by the agreement processing model I proposed, which incorporates the agreement mechanism developed in recent work of the Minimalist Program. The central idea is that automatic feature checking is triggered between D and its c-commanded N, when D carries uninterpretable features. The predictions based on the agreement processing model were supported by the data from L1 processing.

The L2 speakers' overall insensitivity to grammatical agreement violations was interpreted as an indication that automatic syntactic computation was not invoked. Following shallow structure processing hypothesis (SSH), I argued that L2 processing is mainly guided by lexical-semantic and phonetic information, therefore L2 speakers were able to ignore the agreement cues on D. Relating L2 speakers' production with perception, it has been shown that perception does not parallel production; L2 speakers were not sensitive to agreement cues on D in perception, irrespective of their nativelike production with respect to Norwegian DP internal agreement. This observation also points to no L1 transfer in L2 processing. The lack of automatic syntactic computation is therefore not related to the availability of the uninterpretable features on D in L2 speakers' interlanguage grammar. So contrary to Clahsen and Felser (2006), who attribute non-nativelike L2 processing to inadequacy of L2 grammar, I argue instead that L2 processing may be governed by different processing mechanisms. In other words, the full parsing route may not be available to L2 speakers even if they have developed nativelike grammatical representations. In addition, I have argued that the SSH's positioning of morphosyntactic agreement in the morphology side of the morphology-syntax dichotomy might be misguided. Both theoretical considerations and

empirical evidence point to the other direction, i.e. morphosyntactic agreement is more inclined to the syntactic domain.

In brief, the present study points to a difference between L1 and L2 processing. Given the observation that adult L2 speakers remain insensitive to agreement violations in the L2 grammar despite their nativelike grammatical competence, I conclude that the observed L1/L2 difference can be attributed to different processing mechanisms between L1 and L2 speakers. L1 speakers employ the full processing route, whereas L2 speakers use the shallow processing route; the full processing route is probably subject to a critical period and may be no longer available to adult L2 learners.

7 *Conclusions and future directions*

7.1 Summary and conclusion

The aim of this thesis has been to determine whether or not adult second language learners can ultimately achieve linguistic competence that is identical to that of native speakers, not only at the level of grammatical knowledge, but also at the level of processing. This issue has been addressed through examining the production and perception of Norwegian DP internal agreement by three groups of endstate L2 speakers from typologically different backgrounds (Italian/Spanish, English, and Chinese). Specifically, the following three research questions have been asked:

- RQ1: Can L2 speakers attain nativelike success in grammatical competence with respect to Norwegian DP internal agreement? If not, which of the current proposals is most compatible with the L2 data?
- RQ2: Do endstate L2 speakers process Norwegian DP internal agreement the same way as the natives do?
- RQ3: Is L2 speakers' nativelike (or non-nativelike) processing directly related to their L2 grammatical competence? In other words, does perception parallel production?

Results from the on-line DP production task provided an answer to RQ1. It has been found that not all post-critical L2 learners can ultimately develop nativelike grammatical representations that are identical to those of native speakers. While four out of five Romance-speaking learners and three out of six English-speaking learners performed nativelike in the on-line task, the remaining nine speakers (including one Romance-speaking, three English-speaking and all Chinese-speaking learners) fell short of nativelike performance. The two competing SLA models, the FFFH and the FTFA, along with its proposals accounting for NS/NNS divergence (namely, the MSIH and the PTH), have been evaluated in relation to the L2 data. It appears that both models only made the right predictions for a subset of the L2 speakers. The FTFA made the right predictions for the 'nativelike' performers; the FFFH made the right predictions for the 'non-nativelike' performers.

In response to RQ2, results from the auditory naming task indicate that endstate L2 speakers do not process Norwegian DP internal agreement in a nativelike fashion.

While native speakers have been found to show sensitivity to gender/number/definiteness cues on the prime, L2 speakers were generally insensitive to any of the agreement cues. This finding has been interpreted as an indication that L2 processing is less automatic than L1 processing. Moreover, we found evidence pointing to no L1 transfer effects on L2 processing. No gender/number priming effects were found in those L2 groups whose L1s have agreement properties similar to those of the L2. These findings altogether lend support to the SSH, according to which L2 learners underuse structural information when processing the target language.

Turning to RQ3, our findings pertaining to the interrelation between perception and production provided a negative answer to it. That is to say, perception may not parallel production. A comparison as to the processing performance of G1 participants (the so-called ‘nativelike performers’) and G2 participants (the so-called ‘non-nativelike performers’) revealed a remarkably similar processing pattern between the two groups—both are totally insensitive to grammatical agreement violations in the L2 grammar, contra to what has been found in native speakers. The observation suggests that L2 speakers have not attained nativelike processing even after they have developed nativelike grammatical representations.

Summing up, although it remains to be determined to what extent adult L2 speakers are subject to maturational constraints at the level of grammatical knowledge, results from the empirical experiments point to the conclusion that a critical period may affect L2 speakers at the level of processing. For those ‘nativelike performers’, the case may well be that they have attained nativelike grammatical knowledge, but are not able to use it in real-time L2 processing due to unavailability of certain processing mechanisms.

7.2 Contributions

Empirical research can be performed to verify or falsify theories, develop new theories or modify existing ones, and improve research methods. This thesis contributes to all these aims by assessing existing SLA models and processing theories in relation to the L2 data obtained from production and processing experiments, and using the results to propose some refinement of the theories.

a. *Verifying/falsifying theories*: Results from the on-line DP production task revealed that nativelike grammatical knowledge can be acquirable by some L2 speakers, but not by others from the same L1 backgrounds. This within-group variability is not expected by either the FFFH or the FTFA, as both predict uniform performance of an individual L2 group. The results from the production experiment, therefore, are largely inconclusive as to the availability of parameterized uninterpretable features to adult L2 learners. Further investigation of grammatical representations in endstate L2 learners is necessary before any firm conclusions can be drawn.

The data obtained from the ‘non-nativelike performers’ were used to assess the models’ respective proposals (the RDH vs. the MSIH and the PTH) seeking to explain NS/NNS divergence. The evidence for a strong L1 effect on L2 production of agreement errors points to syntactic representational problems in interlanguage grammars of the ‘non-nativelike performers’. Therefore, the RDH is favored over the MSIH and the PTH in explaining L2 learners’ divergent performance.

b. *Modifying existing theories:* L1 and L2 processing performance in the auditory naming task in large part lends support the SSH, according to which L2 processing is restricted to shallow computation, whereas L1 processing typically involves full syntactic representations. However, the theory differentiates between relatively simple morphological rules and the computation of complex syntactic representations. Shallow processing is claimed to be restricted to syntactic processing, whereas for morphological processing, L2 learners are said to be able to employ the same mechanism as native speakers. It is also claimed that morphosyntactic processing, the property under investigation in the current study, falls within the morphological domain (which has also been referred to as ‘local’ domain) of the morphology-syntax dichotomy. However, this claim seems to raise both theoretical and empirical problems. At the theoretical level, the current agreement theory (e.g. Chomsky 1995) assumes that any type of agreement, be it syntactic or morphosyntactic, necessarily involves ‘checking’ of formal features between two constituents sitting in a c-command relationship. Thus, it is not justified to claim that morphosyntactic agreement involves a shallower internal structure than sentences. Moreover, the theory is vague about what is meant by ‘local’ domain. If ‘local’ means two constituents being adjacent to each other, as Clahsen and Felser (2006b) seems to claim, the relationship between D and N is not local at all. Given Julien’s (2005) DP model, there are several intervening heads (e.g. Num, *n*, α) between D and N. Empirically, natively-like processing was not observed. Rather, L2 processing was found to be less automatic than L1 processing, suggesting that a shallow processing route was employed by L2 speakers. Both theoretical considerations and empirical evidence point to the need to put morphosyntactic agreement (at least between D and N) in the domain of syntax, where L2 processing is expected to differ from L1 processing.

Furthermore, the basic assumption of the SSH that L2 processing is different due to inadequacies of L2 grammars is challenged by the non-natively-like processing performance displayed by the ‘natively-like performers’ in the on-line DP production task. If we assume (as is generally believed) that on-line methods are reliable in ‘tapping’ learners’ implicit grammatical knowledge, it follows that the ‘natively-like performers’ have developed natively-like grammatical representations with respect to Norwegian DP internal agreement. The L2 learners’ differential processing performance then can only be attributed to their inability to use this knowledge in real-time L2 processing. This means that Clahsen and Felser’s (2006b:117) second possibility that full parsing fails due to the unavailability of the required parsing mechanism may be proved to be on the right track. If this is so, the basic assumption underlying the SSH has to be revised. Certainly it is still premature to draw any strong conclusions about the interrelation between grammatical representations and grammatical processing in L2 speakers. As has been noted before, little research addresses this issue. The present study is limited with respect to the linguistic phenomenon it looked into and the size of L2 groups it tested on. Further research is needed in order to find out whether the results from this study can be replicated and generalized for other linguistic phenomena.

c. *Developing new theories:* I proposed an agreement processing model in order to explain how feature checking is going on in real-time processing. The model draws on the agreement mechanism proposed in the Minimalist program. The assumption is that syntactic computation takes place during the syntactic formation of an utterance and during its parsing as well. In the case at hand, the uninterpretable features on D will

automatically drive the parsing process, through which a Probe-Goal relation between D and N is established. If D and N have the matching features as required, parsing will be efficient; otherwise, in case there is a feature mismatch, attention will be directed to the mismatched feature, resulting in slowing down in processing. This agreement processing model has been shown to make right predictions for L1 processing. Concordant gender/number/definiteness-marking cases were found to be responded to significantly faster than the discordant counterparts, indicating that feature checking was indeed going on in L1 processing. This processing model is also found to be useful in explaining L2 processing performance. The insignificant RT differences between concordant and discordant cases found in L2 speakers are indicative of absence of automatic computation in L2 processing. In this way, the agreement processing model provides a satisfactory account of how L2 processing differs from L1 processing. In addition, this model is a good illustration of how linguistic theories can be usefully brought to provide an explicit processing framework and to account for results from psycholinguistic experiments.

d. *Improving research methods*: An extension of the concept of ‘linguistic competence’ to let it include both *grammatical knowledge* and *parsing capacities* was taken as the starting point for constructing the relevant experimental situations. As has been noted, the two aspects of linguistic competence typically have been studied in separate realms of L2 research. The former (i.e. grammatical knowledge in L2 learners) has been the major concern in the field of second language acquisition (SLA), and linguistic competence has largely been restricted to grammatical competence. The latter (i.e. L2 processing performance) has been the subject matter of psycholinguistic research. This study has shown that a combination of the approaches in the two disciplines can lead to a better understanding of how adult L2 speakers differ from native speakers. If we take NNS-NS differences as the central goal of L2 research, it is not only crucial to pinpoint in what grammatical aspects adult L2 learners can achieve nativelike grammatical competence, but also important to find out whether they have the ability to use this knowledge in real-time processing. Thus I see the linking of competence-oriented approaches in SLA (particularly using on-line tasks) and psycholinguistic experimentation as a promising research method, which can be used to shed new light on central issues of L2 research.

7.3 Directions for future work

The results of this thesis, its contributions and limitations, point to several possibilities for future research, which bear relevance to theory and methodology.

a. *Identifying endstate L2 learners*: Many researchers (e.g. White 2003; Long 2003; Lardiere 2006a) have pointed out that the best method to identify endstate L2 learners is to use longitudinal data. Due to practical reasons, it was not possible for this study to collect L2 data over a long time span. Instead, a proficiency test and somewhat subjective selection criteria have been used to determine whether the L2 learners were at the endstate. In future work, longitudinal L2 data based on spontaneously produced speech of the L2 speakers will be supplemented. An analysis of the learners’ DP

internal agreement errors will be provided to see whether their performance in naturalistic speech parallels that in the on-line DP production task.

b. *Further testing of the SLA models:* The current results are inconclusive about the availability of novel uninterpretable features to L2 learners. The FFFH and the FTFA are clearly in need of further testing. I would like to highlight the methods used in these tasks which I believe are useful in this regard, that is, use of on-line tasks, comparative L2 groups and analysis of group data as well as individual data. First, data obtained from on-line tasks are more revealing of learners' internalized grammatical knowledge, because the time pressure exerted on learners in online tasks reduces the extent to which learners have access to metalinguistic knowledge. Results from this study have shown that although the L2 speakers appeared to have the same linguistic knowledge about the target language at the metalinguistic level, their performance in the on-line task indicated they had different implicit grammatical representations. So in further work, where it is possible, on-line tasks will be designed to tap learners' grammatical competence. Second, comparative L2 groups can be crucial in assessing different SLA theories. To take just one example from the present study, the PTH would predict that Chinese-speaking learners' production of /t/ will be depressed in contrast with /e/, and that English-speaking learners should be accurate in their production of both inflections. It was found that the data from the Chinese group were consistent with the predictions of the PTH. But a comparison of the data from the English group showed that English-speaking learners exhibited a similar pattern to the Chinese group. Thus the PTH turned out to be untenable. If we had only used the Chinese group, a misleading conclusion would probably be drawn. Lastly, I would like to emphasize the need to analyze individual data along with the group data. The significance of analyzing individual data was pointed out by White (2003):

Since the claim that the interlanguage grammar is (or is not) UG-constrained is a claim about individual linguistic competence, it is crucial to determine what is going on at the individual level (White 2003:55).

This study has shown that learners from the same L1 backgrounds may not display uniform performance. Thus it is necessary to analyze individual data in order to find the differences in learners' idiolects.

c. *Further testing of the agreement processing model and the SSH:* The agreement processing model proposed in this study has been shown to make the right predictions for the L1 processing of local mismatches in terms of gender, number and definiteness. It also served to explain findings obtained from L2 processing. Clearly, the model does not constrain itself to explaining local-domain feature checking. Future research involving non-local domain feature mismatches in L1 and L2 processing is necessary to further test the agreement processing model, as well as the SSH. In this aspect, Norwegian also provides a good testing point, as it also exhibits gender and number agreement between subjects and predicative adjectives, as shown in (7.1):

- (7.1) a. Hus-et er fin-t.
 house-NEUT.SG.DEF is nice-NEUT.SG.
 'The house is nice.'

- b. Bil-en er fin.
car-COM.SG.DEF is nice-COM.SG.
'The car is nice'
- c. Hus-ene/Bil-ene er fin-e
house-PL.DEF/car-PL.DEF is nice-PL
'The houses/cars are nice.'
- d. Biler er fine
car-PL.INDEF is nice-PL
'The cars are nice.'

In light of the agreement processing model, I shall expect similar contrasting performance between L1 and L2 speakers. Furthermore, it will be of interest to compare the performance between early bilinguals and late bilinguals, in order to assess the age factor in L2 processing. If early bilinguals and late bilinguals are shown to have the same grammatical representations in terms of the aspects of the grammar being tested, but their behavior is fundamentally different in processing, it will provide more revealing evidence for the hypothesis that the full parsing route is subject to a critical period. In addition, the question raised by the present study will also be addressed in future work. As has been mentioned earlier, the role of L1 transfer in L2 processing may in some way be influenced by the type of the task, so in future work different techniques (for example, EEG and auditory naming) will be used, and different tasks (for example, reading the feature matched and mismatched expressions in sentence context) will be included. Comparing results obtained from different learner groups, different tasks, different techniques, and different linguistic domains will hopefully lead us to a better understanding about the nature of L1 and L2 processing.

As a concluding note, I call for more L2 research addressing the intrinsic relationship between grammatical competence and grammatical processing. I would like to reemphasize that linking competence-oriented approaches in SLA and psycholinguistic experimentation can be useful in this regard. I believe that such research can shed new light on maturational constraints in second language acquisition, thus contributing to our understanding about the nature of the language faculty.

References

- Adger, D. (2003) *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.
- Akhutina, T., Kurgansky, A. Polinsky, M. and Bates, E. (1999) Processing of grammatical gender in three-gender system: Experimental evidence from Russian. *Journal of Psycholinguistic Research* 28, 695-713.
- Anderson, M. (2004) The acquisition of compositional definiteness in Norwegian. Unpublished doctoral dissertation, University of Tromsø.
- Andonova, E., D'Amico, S., Devescovi, A. and Bates, E. (2004) Gender and lexical access in Bulgarian. *Perception and Psychophysics* 66, 496-507.
- Askedal, J.O. (1994) Norwegian. In E. König and J. van der Auwera (eds.) *The Germanic Languages* (pp.219-270). London and New York: Routledge.
- Baker, M. C. (2003) *Lexical Categories*. Cambridge: Cambridge University Press.
- Bates, E., Devescovi, A., Pizzamiglio, L., D' Amico, S. and Hernandez, A. (1995) Gender and lexical access in Italian. *Perception and Psychophysics* 57, 847-862.
- Bates, E., Devescovi, A., Hernandez, A., and Pizzamiglio, L. (1996) Gender priming in Italian. *Perception & Psycholinguistics* 58, 992-1004.
- Bates, E. and Liu, H. (1997) Cued shadowing. In F. Grosjean and U. Frauenfelder (eds.) *A Guide to Spoken Word Recognition Paradigms* (pp. 577-583). Hove, U.K.: Psychology Press.
- Boeckx, C. (2006) *Linguistics Minimalism: Origins, Concepts, Methods and Aims*. Oxford: Oxford University Press.
- Boersma, P. and Weenink, D. (2006) Praat: doing phonetics by computer (Version 4.4.11) [Computer program]. Retrieved February 23, 2006, from <http://www.praat.org/>.
- Bohnacker, U. (2004) Nominal phrases. In G. Josefsson, C. Platzack and G. Håkansson (eds.) *The Acquisition of Swedish Grammar*. Amsterdam/Philadelphia: John Benjamins.
- Brown, C. (2000) The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (ed.) *Second Language Acquisition and Linguistic Theory* (pp. 4-63). Malden, MA: Blackwell.
- Carroll, S. (1989) Second language acquisition and the computational paradigm. *Language Learning* 39, 535-594.
- Carstens, V.M. (2000) Concord in minimalist theory. *Linguistic Inquiry* 31, 319-355.

- Chen, L., and R. Sybesma. (1999) Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30:509-542.
- Chierchia, G. (1998) Reference to kinds across languages. *Natural Language Semantics* 6, 399-405.
- Chomsky, N. (1980) *Rules and Representations*. Oxford: Blackwell.
- Chomsky, N. (1986) *Knowledge of Language: its Nature, Origin and Use*. Praeger.
- Chomsky, N. (1995) *The Minimalist Program*. Cambridge, Mass.: The MIT Press.
- Chomsky, N. (1999) Derivation by phase. *MIT occasional Papers in Linguistics* 18. Cambridge, Mass.: MITWPL.
- Chomsky, N. (2000) Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka, (eds.) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89-155. Cambridge, Mass.: The MIT Press.
- Chomsky, N. (2001) *Beyond Explanatory Adequacy*. Cambridge, MA: MIT Press.
- Chomsky, N. (forthcoming) On phases. In R. Freidin, C. Otero and M.-I. Zubizarreta (eds.) *Foundational Issues in Linguistic Theory*. Cambridge, MA: MIT Press.
- Clahsen, H. and Felser C. (2006a) Grammatical processing in language learners. *Applied Psycholinguistics* 27, 3-42.
- Clahsen, H. and Felser C. (2006b) Continuity and shallow structures in language processing. *Applied Psycholinguistics* 27, 107-126.
- Corbett, G. (1991) *Gender*. Cambridge: Cambridge University Press.
- Correa, L.-S., Augusto, M. and Ferrari-Neto, J. (2005). The early processing of number agreement in the DP: evidence from the acquisition of Brazilian Portuguese. Paper presented at the 30th Boston University Conference on Language Development, 4-6 November.
- DeKeyser, R. M. (2000) The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22, 493-533.
- Delsing, L.-O. (1993) The internal structure of noun phrases in the Scandinavian languages. Unpublished doctoral dissertation, Lund University.
- Dörnyei, Z. and Skehan, P. (2003) Individual differences in second language learning. In C. Doughty and M. Long (eds.) *Handbook of Second Language Acquisition* (pp.589-630). Oxford: Blackwell.

- Eubank, L. and Grace, S. T. (1998) V-to-I and inflection in non-native grammars. In M.-L. Beck (ed.), *Morphology and its Interface in L2 Knowledge* (pp.69-88). Amsterdam: John Benjamins.
- Ellis, R. (2005) Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition* 27, 141-172.
- Foucart, A. and Frenck-Mestre, C. (2004) *Processing of grammatical gender information in French as first and second language*. Poster presented at AMLaP, Aix-en Provence.
- Franceschina, F. (2001) Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, 17, 213-247.
- Franck, J., Lassi, G., Grauenfelder, H-U. and Rizzi, L. (2006) Agreement and movement: A syntactic analysis of attraction. *Cognition* 101,173-216.
- Gillon-Dowens, M., Barber, H., Bergara, M. and Carreiras, M. (2004) *Does practice make perfect? An ERP study of morphosyntactic processing in highly proficient English-Spanish late bilinguals*. Poster presented at AMLaP, Aix-en Provence.
- Goad, H., White, L. and Steele, J. (2003) Missing inflection in L2 acquisition: Defective syntax or L1-constrained prosodic representation? *Canadian Journal of Linguistics* 48, 243-263.
- Goad, H. and White, L. (2004) Ultimate attainment of L2 inflection: Effects of L1 prosodic structure. In S. Foster-Cohen, M. Sharwood, A. Sorace and M. Ota (eds.) *Eurosla Yearbook 4*. Amsterdam/Philadelphia: John Benjamins.
- Goad, H. and White, L. (2006) Ultimate attainment in interlanguage grammars: A prosodic approach. *Second Language Research* 22, 243-268.
- Grosjean, F., Dommergues, J.-Y., Cornu, E., Guillelmon, D. and Besson, C. (1994) The gender-marking effect in spoken word recognition. *Perception and Psychophysics* 56, 590-598.
- Guillelmon, D. and Grosjean, F. (2001) The gender marking effect in spoken word recognition: The case of bilinguals. *Memory and Cognition* 29, 503-511.
- Hagoort, P. and Brown, C. (1999) Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of psycholinguistic Research* 28, 715-728.
- Hahne, A. and Friederici, A. D. (2001) Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition* 4, 123-141.

- Hahne, A., Müller, J. and Clahsen, H. (2003) Second language learners' processing of inflected words: Behavioral and ERP evidence for storage and decomposition. *Essex Research Report in Linguistics* 45, 1-42.
- Halle, M. (1997) Distributed morphology: Impoverishment and fission. *MIT Working Papers in Linguistics* 30, 425-449.
- Halle, K. and Keyser, S.J. (2002) *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: MIT Press.
- Halle, K. and Marantz, A. (1993) Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyser (eds.) *The View from Building 20* (pp.111-176). Cambridge: MIT Press.
- Hawkins, R. (2001a) The theoretical significance of Universal Grammar in second language acquisition. *Second Language Research* 17, 345-367.
- Hawkins, R. (2001b) *Second Language Syntax: A Generative Introduction*. Oxford: Blackwell.
- Hawkins, R. (2004) Revisiting *wh*-movement: The availability of an uninterpretable [wh] feature in interlanguage grammars. *Cascadilla Proceedings Project*. Somerville, MA: Cascadilla Press.
- Hawkins, R., Al-Eid, S., Almahboob, I., Hu, Y., Jaensch, C., Jeon, Y., Leung, I., Ranong, S., Ortega, M., Polentas, A., Rezai, M., Roehr, K., Sarko, G., Snape, N. and Sarantis K. (2006) Knowledge of English resultatives by L2 speakers— evidence from UG-constrained interlanguage grammars. Paper presented at the 16th Conference of the European Second Language (Eurosla-16), Antalya, Turkey, September.
- Hawkins, R. and Chan, C.Y.-H. (1997) The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research* 13, 187-226.
- Hawkins, R. and Franceschina, F. (2003) Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. In J. Paradis and P. Prévost (eds.) *The Acquisition of French in Different Contexts* (pp.175-205). Amsterdam /Philadelphia: John Benjamins.
- Hawkins, R. and Hattori, H. (2006) Interpretation of English multiple *wh*-questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research* 22, 269-301.
- Hawkins, R. and Liszka, S. (2003) Locating the source of defective past tense marking in advanced L2 English speakers. In R. van Hout, A. Hulk, F. Kuiken and R. Towell (eds.) *The Interface between Syntax and Lexicon in Second Language Acquisition* (pp.21-44). Amsterdam/Philadelphia: John Benjamins.

- Horstein, N. (2001) *Move! A Minimalist Approach to Construal*. Oxford: Blackwell.
- Huang, C-T.J., Li, A., and Li, Y. (to appear) *The Syntax of Chinese*. Cambridge: Cambridge University Press.
- Hyltenstam, K. and Abrahamsson, N. (2003) Maturational constraints in SLA. In C. Doughty and M. Long (eds.) *Handbook of Second Language Acquisition* (pp. 539-588). Oxford: Blackwell.
- Josefsson, G. and Platzack, C. (2004) Introduction. In G. Josefsson, C. Platzack and G. Håkansson (eds.) *The Acquisition of Swedish Grammar*. (pp. 1-22). Amsterdam/Philadelphia: John Benjamins.
- Julien, M. (2003) Double definiteness in Scandinavian. *Nordlyd* 31, 230-244.
- Julien, M. (2005) *Nominal Phrases from a Scandinavian Perspective*. Amsterdam: John Benjamins.
- Kanno, K. (1997) The acquisition of null and overt pronominals in Japanese by English speakers. *Second Language Research* 13, 265-287.
- Kanno, K. (1998) The stability of UG principles in second language acquisition: Evidence from Japanese. *Linguistics* 36, 1125-1146.
- Kester, E.-P. (1996) The nature of adjectival inflection. Unpublished doctoral dissertation, Utrecht University.
- Klapka, L. (2002) L'accord du genre en français québécois au 19ième siècle. Paper presented at the Annual Meeting of the Canadian Linguistics Association, Toronto, May.
- Kurafuji, T. (2004) Plural morphemes, definiteness, and the notion of semantic parameter. *Language and Linguistics* 5, 211-242.
- Kuroda, S-Y. (1988) Whether we agree or not: A comparative syntax of English and Japanese. In W. Poser (ed.) *Papers from the Second International Workshop on Japanese Syntax* (pp. 103-143). Stanford: CSLI.
- Lardiere, D. (1998a) Case and tense in the 'fossilized' steady state. *Second Language Research* 14, 1-26.
- Lardiere, D. (1998b) Dissociating syntax from morphology in a divergent end-state grammar. *Second Language Research* 14, 359-375.
- Lardiere, D. (2000) Mapping features to forms in SLA. In Archibald, J. (ed.) *Second Language Acquisition and Linguistic Theory* (pp. 102-129). Oxford: Blackwell.

- Lardiere, D. (2006a) *Ultimate Attainment in Second Language Acquisition: A Case Study*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Lardiere, D. (2006b) Acquiring (or assembling) functional categories in second language acquisition. Paper presented at the 2nd Generative Approach to Language Acquisition North America Conference (GALANA-2), Montreal, August.
- Lardiere, D. (2006c) Introduction. *Second Language Research* 22, 239-242.
- Lenneberg, E. (1967) *Biological Foundations of Language*. New York: John Wiley.
- Leung, Y.-K. I. (1998) Transfer between interlanguages. In A. Greenhill, M. Hughes, H. Littlefield and H. Walsh (eds.) *Proceedings of the 22nd Boston University Conference on Language Development* (pp. 477-487). Somerville, MA: Cascadilla Press.
- Li, C. N. and Thompson, S. (1981) *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Ijic, R. (1994) Qualification in Mandarin Chinese: Two markers of plurality. *Linguistics* 32, 91-116.
- Long, M. (1997) Fossilization: Rogor mortis in living linguist systems? Plenary address to the Eurosla Conference, Barcelona, May.
- Long, M. (2003) Stabilization and fossilization in interlanguage development. In C. Doughty and M. Long (eds.) *Handbook of Second Language Acquisition* (pp.487-535). Oxford: Blackwell.
- Lozano, C. (2003) Universal Grammar and focus constraints: The acquisition of pronouns and word order in non-native Spanish. Unpublished doctoral dissertation, University of Essex.
- Lü, Shuxiang. (1980) *Xiandai Hanyu Babai Ci* (800 Words in Modern Chinese). Beijing: Commercial Press.
- Lück, M., Hahne, A., Friederici, A. and Clahsen, H. (2001) Developing brain potentials in children: An ERP study of German noun plurals. Paper presented at 26th Boston University Conference on Language Development, November.
- Lyons, C. (1998) *Definiteness*. Cambridge: Cambridge University Press.
- Marinis, T., Roberts, L., Felser, C. and Clahsen, H. (2005) Gaps in second language sentence processing. *Studies in Second Language Acquisition* 27, 53-78.
- Penke, M., Weyerts, H., Gross, M., Zander, E., Münte, T. and Clahsen, H. (1997) How the brain processes complex words: An event-related potential study of German verb inflections. *Cognitive Brain Research* 6, 37-52.

- Pérez-Leroux, A-T. and Glass, W. (1999) Null anaphora in Spanish second language acquisition: Probabilistic versus generative approaches. *Second Language Research* 15, 220-249.
- Prévost, P. and White, L. (2000) Missing Surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16, 103-33.
- Radford, A. (2004) *Minimalist Syntax: Exploring the Structure of English*. Cambridge: Cambridge University Press.
- Ritter, N. (2002) Introduction. *The Linguistic Review* 17, 1-7.
- Rizzi, L. (1982) *Issues in Italian Syntax*. Dordrecht: Foris.
- Robertson, D. (2000) Variability in the use of the English article system by Chinese learners of English. *Second Language Research* 16, 35-172.
- Robertson, D. and Sorace, A. (1999) Losing the V2 constraint. In E. Klein and G. Martohardjono (eds.) *The Development of Second Language Grammars: A Generative Approach* (pp. 317-61). Amsterdam/Philadelphia: John Benjamins.
- Sabourin, L. (2003) Grammatical gender and second language processing: An ERP study. Unpublished doctoral dissertation, University of Groningen.
- Schwartz, B. D. and Sprouse, R. (1994) Word order and nominative case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In T. Hoekstra and B. D. Schwartz (eds.) *Language Acquisition Studies in Generative Grammar* (pp. 317-368). Amsterdam/Philadelphia: John Benjamins.
- Schwartz, B. D. and Sprouse, R. (1996) L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12, 40-72.
- Smith, N. and Tsimpili, I.-M. (1995) *The Mind of a Savant: Language Learning and Modularity*. Oxford: Blackwell.
- Svenonius, P. (1993) Selection, adjunction, and concord in the DP. *Studia Linguistica* 47, 198-220.
- Taraldsen, K. T. (1990) D-projections and N-projections in Norwegian. In J. Mascaró and M. Nespór (eds.) *Grammar in Progress*, Foris: Dordrecht.
- The Oslo Corpus of Tagged Norwegian Texts. <http://www.tekstlab.uio.no/>
- Thomas, R. (2002) Development of the concept of “the poverty of the stimulus.” *The Linguistic Review* 17, 51-71.

- Tokowicz, N. and MacWhinney, B. (2005) Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition* 27, 173-204.
- Travis, L. (1984) Parameters and the effects of word order variation. Unpublished doctoral dissertation, MIT.
- Tsimpli, I.-M. (2003) Features in language development. Plenary paper presented at EUROSLA, Edinburgh, UK., September.
- Tsimpli, I.-M. and Roussou, A. (1991) Parameter-resetting in L2? *UCL Working Papers in Linguistics* 3, 149-170.
- Ullman, M. (2004) Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92, 231-270.
- Vainikka, A. and Young-Scholten, M. (1994) Direct access to X'-theory: Evidence from Turkish and Korean adults learning German. In T. Hoekstra and B.D. Schwartz (eds.) *Language Acquisition Studies in Generative Grammar* (pp. 265-316). Amsterdam: Benjamins.
- van Berkum, J. (1996) *The Psycholinguistics of Grammatical Gender*. Nijmegen: Nijmegen University Press.
- Vangsnes, Ø. A. (1999) The identification of functional architecture. Unpublished doctoral dissertation, University of Bergen.
- Weber-Fox, C. and Neville, H. (1996) Sensitive periods differentiate processing of open- and closed class words: An ERP study of bilinguals. *Journal of Speech, Language and Hearing Research* 44, 1338-1353.
- Weyerts, H., Penke, M., Dohrn, U., Clahsen, H. and Münte, T. (1997) Brain potentials indicate differences between regular and irregular German plurals. *NeuroReport* 8, 957-962.
- White, L. (1989) *Universal Grammar and Second Language Acquisition*. Amsterdam: Benjamins.
- White, L. (2000) Second language acquisition: From initial to final state. In J. Archibalds (ed.) *Second Language Acquisition and Linguistic Theory* (pp.130-155). Oxford: Blackwell.
- White, L. (2002) Morphological variability in endstate L2 grammars: the question of L1 influence. In B. Skarabela, S. Fish, and A.H.-J. Do (eds.) *Proceedings of the 26th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

White, L. (2003a) *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

White, L. (2003b) Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition* 6, 129-141.

White, L., Valenzuela, E., Kozłowska-Macgregor, M. and Leung, Y.-K. I. (2004) Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics* 25, 105-133.

Xu, L. and Langendoen, T. (1985) Topic structures in Chinese. *Language* 61, 1-27.

Yang, H. (2004) Classifiers and plurality in Mandarin NP. Paper presented at GLOW 27, Thessaloniki, April.

Yoon, J. (1999) Notes on 'Derivation by Phase'. Draft, downloadable from <http://www.punksinscience.org/kleanthes/courses/UK03S/OSPS/Yoon.pdf>.

Appendix

1. Norwegian Proficiency Test

Luketest: Sett in ett ord hver luke.

ALKOHOLPOLITIKK I NORDEN

Mange utlendinger (1) kommer til Norge for første gang, reagerer (2) at det er dyrt og til dels vanskelig å få tak (3) alkohol. Øl kan kjøpes i (4) dagligvarebutikker, mens vin og brennevin (5) kjøpes på Vinmonopolet. Også Sverige og Finland har et (6) alkoholmonopol, mens danskene kan kjøpe (7)vin og brennevin i dagligvarebutikker. (8) andre ord er alkohol lettere tilgjengelig i Danmark enn i de andre nordiske (9)..... Norge, Island og Sverige har (10) laveste alkoholforbruket i Europa med ca. fem liter ren alkohol per innbygger per år. Danmark og Finland (11) på europeisk mellomnivå med over ni liter. Irland har (12)..... største alkoholkonsumet i Europa med nesten 12 liter (13) hver innbygger per år.

Siden Norge (14) er medlem av EU, kan vi (15)en viss grad selv bestemme vår (16) alkoholpolitikk. I Norge er prisen på alkohol (17) enn i våre nordiske naboland som er med i EU. Høye priser skyldes høye avgifter til Staten. Myndighetene forsvarer (18) høye avgiftene med at alkohol er skadelig for helsa. (19) våre nordiske naboer kan ta med (20)nesten ubegrensede mengder alkohol fra andre EU-land, er kvotene i Norge en liter sprit og en liter vin per person over 18 år.

Mange synes at dette er altfor (21)og tar med seg litt ekstra. Det er nesten blitt (22)..... folkesport her i landet å lure tollmyndighetene. I (23) smugles det store mengder sprit til Norge fra Øst- og Sør-Europa. Noe av (24) spriten har vist seg å være giftig, og det har ført (25) mange dødsfall i Norge.

I Danmark forklarer man de lave (26) med at danskene har felles grense med Tyskland. (27) mange reiser over grensa for å kjøpe drikkevarer, kjøper

de gjerne med seg andre varer også, og det går ut over landets eget næringsliv. Finland fikk også erfare dette (28)..... nabolandet Estland ble medlem av EU våren 2004. Det er kort avstand (29) Finland og Estland, og de (30) myndighetene fryktet for at det ville bli en stadig større import fra Estland. (31) halverte finnene prisen på sprit for å unngå store konsekvenser for (32) eget næringsliv. Alkoholprisene i Norge og Sverige er temmelig like. Sverige er (33) medlem av EU, og svenskene kan importere store mengder alkohol tollfritt. Med nært naboskap til Danmark og Finland vurderer svenskene nå å sette (34) alkoholprisene til dansk og finsk nivå. Norge har en lang felles grense med Sverige, og prisreduksjonen i vårt naboland kan få (35) konsekvenser her til lands, tror mange norske politikere. De frykter også mer ulovlig import (36)alkoholholdige drikkevarer. Hvis svenskene reduserer prisene (37)40%, kommer de på linje med Danmark, men danskene har planer (38) å senke prisene ytterligere for å komme på linje med Øst- og Sør-Europa. Dette er (39)dominoeffekt hvor den ene påvirkes av den (40), og vi vet ikke hvor denne utviklingen vil ende.

Det fins allerede lange tradisjoner i Norge for (41) folk i grenseområdene til Sverige handler billige varer i nabolandet. Mange trøndere tar billigbuss til (42) svenske bygda Storlien flere ganger (43) måneden for å kjøpe sprit og tobakk. Statistikk (44) at nordmenn handler for mange milliarder kroner i Sverige og andre naboland, og (45) tallet er mye høyere enn de beløpene våre naboer legger (46)..... i våre (47) butikker. Dette er (48)uheldig utvikling for Norge. Staten tjener på de høye alkoholavgiftene, men samtidig (49) samfunnet mange arbeidsplasser innenfor produksjons- og handelsnæringene. Økonomi er en viktig årsak til de høye avgiftene i Norge. En (50) årsak er sosial- og helsepolitikken.

Key: 1. som; 2. på; 3. i; 4. vanlige; 5. må; 6.likende; 7. både; 8. Med; 9. landene; 10. det; 11. ligger; 12. det; 13. for; 14. ikke; 15. til; 16.egen; 17. høyere; 18.de; 19. Mens; 20. seg; 21. lite; 22. en; 23. tillegg; 24. denne; 25. til; 26. prisene; 27. Hvis; 28.da; 29. mellom; 30. finske; 31. Derfor; 32. sitt; 33. imidlertid; 34. ned; 35. store; 36. av; 37.med; 38.om; 39. en; 40. andre; 41. at; 42. den; 43. i; 44. viser; 45. dette; 46. igjen; 47. norske; 48.en; 49. mister; 50. annen.

2. Gender assignment task

Eksampler: 1. <i>et en</i> dag			2. <i>et en</i> eple	
<i>et en</i>		sak	<i>et en</i>	område
<i>et en</i>		by	<i>et en</i>	sted
<i>et en</i>		vei	<i>et en</i>	problem
<i>et en</i>		løsning	<i>et en</i>	bil
<i>et en</i>		liv	<i>et en</i>	tema
<i>et en</i>		gutt	<i>et en</i>	inntrykk
<i>et en</i>		bank	<i>et en</i>	våpen
<i>et en</i>		ball	<i>et en</i>	ønske

<i>et en</i>	gave	<i>et en</i>	symbol
<i>et en</i>	mening	<i>et en</i>	bygg
<i>et en</i>	tilfelle	<i>et en</i>	avtale
<i>et en</i>	hus	<i>et en</i>	brann
<i>et en</i>	rom	<i>et en</i>	venn
<i>et en</i>	salg	<i>et en</i>	fordel
<i>et en</i>	kurs	<i>et en</i>	krise
<i>et en</i>	brev	<i>et en</i>	debatt
<i>et en</i>	artikkel	<i>et en</i>	system
<i>et en</i>	firma	<i>et en</i>	lån
<i>et en</i>	replikk	<i>et en</i>	feil
<i>et en</i>	forfatter	<i>et en</i>	forslag
<i>et en</i>	leilighet	<i>et en</i>	tilbud
<i>et en</i>	fugl	<i>et en</i>	tall
<i>et en</i>	kirke	<i>et en</i>	sykehus
<i>et en</i>	undersøkelse	<i>et en</i>	skade
<i>et en</i>	land	<i>et en</i>	skole
<i>et en</i>	folk	<i>et en</i>	svar
<i>et en</i>	arbeid	<i>et en</i>	bord
<i>et en</i>	navn	<i>et en</i>	produkt
<i>et en</i>	valg	<i>et en</i>	medlem
<i>et en</i>	lege	<i>et en</i>	kamp
<i>et en</i>	forhold	<i>et en</i>	oppdragelse
<i>et en</i>	kultur	<i>et en</i>	sykkel

3. Instructions for the on-line DP production task

Rettledning: Du vil få høre 60 korte setninger. Når du hører en setning, vil du samtidig få se følgende på skjermen: et substantiv som har forekommet i setningen, og et adjektiv, som står i parenteser. Gjør om setningen, slik at den riktige formen av adjektivet står til substantivet, og si deretter den nye setningen høyt inn i mikrofonen. Selv om du ikke husker hele setningen, prøv å si så mye av den som du kan. Det er bare fem sekunder mellom hver setning, så du må svare raskt. Her er tre eksempler du kan øve deg på:

Det første eksemplet:

Du vil få høre: *Det sitter en **student** i kantina.*

Du vil få se dette på skjermen: { student (lys) }

Forventet respons: *Det sitter en lys **student** i kantina.*

Det andre eksemplet:

Du vil få høre: *Han har sett **filmen**.*

Du vil få se dette på skjermen: { film (morsom) }

Forventet respons: *Han har sett den morsomme **filmen**.*

Øv på det tredje eksemplet for deg selv:

Legg merke til at substantivets form i den nye setningen skal være den samme som i den oppgitte setningen.

4. Short sentences stimuli in the on-line DP production task

1. Han hørte en *replikk*. (lang)
2. Det er interessant å lese *aviser*. (fri)
3. Jeg liker ikke *vintrene* her. (mørk)
4. *Arbeidet* ødela ryggen hans til slutt. (lang)
5. Han slo seg ned i et *boligområde*. (tom)
6. Hun forsto ikke *koden*. (enkel)
7. Han er en *forfatter*. (aktiv)
8. Trondheim var en av *byene* på 1000-tallet. (viktig)
9. Han vil rive deler av et *bygg*. (blå)
10. Han fikk høre *cd-en* hennes. (dyr)
11. De ble påvirket av *kulturer*. (annen)
12. Alle reagerte på *boligprisene*. (høy)
13. På kaffen satt det fire *jenter*. (flink)
14. De bodde in en *leilighet*. (varm)
15. Han kjøpte *genseren*. (gul)
16. *Flertallet* støtter kongen. (klar)
17. De forstår ikke *språket*. (ny)
18. Kongehuset blir sett på som et *symbol*. (nasjonal)
19. Fossekallen er en *fugl*. (pen)
20. I Oslo bor det mange *mennesker*. (populær)
21. Hun vokste opp på et *sted*. (ren)
22. Mange av *symbolene* forandrer seg. (sosial)
23. I *bondesamfunnet* var familien en sentral faktor. (god)
24. Nå blir *saken* undersøkt. (sentral)
25. Han likte ikke *demokratiet*. (spesiell)
26. På Internett finner man *nyhetene* i kortversjon. (frisk)
27. Jeg har et *problem*. (tøff)
28. Det ligger en *kirke* i sentrum. (grei)
29. Bunadene forteller om *forskjeller*. (lokal)
30. Jeg fikk låne *sykkelen* hans. (trygg)
31. Hun hatet *været*. (kald)
32. Har du sett *programmene*? (morsom)
33. De flyttet til et *land*. (nær)
34. Alle ønsker å gi barna en *oppdragelse*. (effektiv)
35. Jeg møtte fem *turister*. (fin)
36. Jeg ble lei av å høre på *diskusjonen*. (sterk)
37. De har en *kultur*. (rik)
38. Han har mange *bøker*. (tung)
39. Hun hadde vært på toppen av *fjellet*. (vill)
40. Jeg kan ikke akseptere *oppførselen* hans. (åpen)
41. Vikingene kom fra alle *landene*. (nordisk)

42. Et *tema* er kloning av mennesker. (annen)
 43. *Forfatteren* Henrik Ibsen ble født i Skien. (stor)
 44. EU står overfor *oppgaver*. (vanskelig)
 45. Hun liker ikke *laget*. (grønn)
 46. Det var en *fest*. (formell)
 47. Han er fornøyd med *forholdene* i Norge. (generell)
 48. Det gjorde et *inntrykk* på ham. (svak)
 49. Samfunnet prøver å dempe *forskjellene* mellom kjønnene. (dyp)
 50. Det står en *artikkel* i avisen. (aktuell)
 51. Han tar med seg et *våpen*. (gammel)
 52. *Svaret* irriterte meg. (lav)
 53. Nordmenn regnes som *turgåere*. (ivrig)
 54. Hun leste *annonsen* en gang til. (gal)
 55. Vi utførte en *undersøkelse*. (offisiell)
 56. Han hadde et *ønske* om å bli dikter. (enorm)
 57. Jeg har lest *artikkelen* i avisen. (liten)
 58. *Klimaet* er ganske bra. (varm)
 59. Det sitter noen *studenter* i kantina. (høy)
 60. Jeg er ivrig å bruke *ordene*. (positiv)

5. Non-target-like DPs produced by the individual participants

R1 (6)	R4 (6)	R5 (6)
19 et blå bygg	6. den enkel kode	3. mørke vintrene
22. de sosiale symboler	16. klare flertall	6. enkel kode
33. et nær land	18. en nasjonal symbol	19. en blå bygg
42. den andre temaet	37. sterke diskusonen	35. fine turistene
56. et enorme ønske	42. en annen tema	57. den liten artikkel
60. positive ordene	57. den liten artikkelen	60. positive ordene
R2 (1)	E1 NL (1)	E4 SP (0)
41. nordiske stedene	22. sosiale symbolene	E6 RG (0)
R3 (23)	E5 (26)	C3 (25)
2. fri avisene	3. mørke vintrene	1. en langt replikk
3. mørke vintren	5. en tom område	6. den enkelt koden
4. lang arbeid	6. enkel koden	8. en viktig byen
6. enkel kode	8. en viktige byen	10. dyrt cden
10. dyre cden	10. dyre cden	11. annen kultur
11. annen kulturer	11. annen kultur	12. den høyt bligprisene
14. en varmt leilighet	12. høye prisene	17. de nye språket
17. nye språket	13. flinke jentene	19. det blått bygg
18. en nasjonal symbol	15. den gule genser	22. de sosial symbolene
25. spesiell demokratiet	16. en klare flertallet	23. en god bondesamfunnet
27. et tøff problem	19. en blå bygg	24. sentral saken
30. trygge sykkel	21. en ren sted	26. friske nyheten
31. kalde været	23. god bondesamfunnet	29. det lokal forskjeller

32. morsomme programene	25. spesielt demokrati	31. det kaldt været
39. vill fjellet	26. frisk nyhet	36. sterke diskusjonen
40. åpen oppførselen	27. et tøff problem	39. den ville fjellet
41. de norske stedet	31. kalde været	41. nordiske stedene
	32. morsome	
43. store forfatteren	programene	43. den stor forfatteren
45. grønne laget	33. et nær land	45. den grønne laget
48. svak inntrykk	45. grønne laget	47. general forholdene
49. dyp forskjell	49. dype forskjellene	52. det lav svaret
52. lav svaret	51. en gammel våpen	56. en enorm ønske
56. en enorm ønske	52. lav svar	57. den liten artikkelen
	56. en enorm ønske	58. den normal klimaet
	57. den liten artikkel	59. høy studentene
	58. den normal klimaet	
E2 (22)	C1 (30)	C4 (34)
2. fri aviser	2. fri aviser	1. fri aviser
3. mørke vintrene	3. det mørke vinter	3. den mørke vintrene
4. lang arbeidet	4. lang arbeidet	4. den lang arbeidet
6. en enkel koden	6. enkelt koden	5. et tom boligområde
10. dyre cden	8. den viktig byen	7. et aktiv forfatter
12. høye prisene	10. den dyre cd	8. viktig byen
16. klar flertall	11. en annen kulturer	11. et annen kultur
17. nye språket	12. høy prisen	12. den høye prisene
22. sosiale symbolene	17. den nye språket	14. et varm leilighet
23. en god		
bondesamfunn	19. det blålig bygg	16. klare flertallet
24. sentral saken	21. et ren sted	17. den nytt språket
25. spesielt demokratiet	23. gode bodesamfunnet	18. et nasjonal symbol
30 et trygg sykkel	25. spesielt demokratiet	19. et blå bygg
36. sterke diskusjonen	28. et grei kirke	21. et ren sted
		23. den gode
37. et rikt kultur	29. de lokale forskjeller	bondesamfunnet
39. vill fjellet	31. den kalde været	27. et tøff problem
	34. et effektivt	
43. stor forfatteren	oppdragelse	28. et grei kirke
45. grønne laget	39. de ville fjellet	29. den lokale forskjeller
		32. den morsomme
47. generelle forhold	40. den åpne oppførsel	programene
48. et svak inntrykk	41. nordiske stedene	33. et nær land
50. en aktuelt artikkel	43. stor forfatteren	34. et effektiv oppdragelse
	44. de vanskelige	
60. positive ordene	oppgaver	39. et ville fjellet
	45. den grønne laget	40. den åpen oppførsel
	47. generell forhold	42. et annen tema
		44. den vanskelig
C5 (29)	48. et svak intrykk	oppgavene
1. de fri avisere	51. en gammelt våpen	45. det grønn laget
3. en mørke vinter	52. den lave svaret	47. det generale forholdene
5. et tom boligområde	56. et enorm ønske	48. et svak inntrykk
9. den pene fuglet	58. den normal klimaet	49. det dype forskjellene
10. de dyre cden	60. de positiv ordene	50. et aktuelle artikkel

11. de andre kulturer		51. et gammel våpen
13. de flinke jenter		55. et offisiell undersøkelse
15. de gull genserne	E3 (23)	57. den liten artikkelen
18. en nasjonal symbol	2. den fri avisen	60. den positive ordene
19. det blå bygg	3. mørke vintrene	
	5. en tom boligområde -	
21. de rene sted	et	C2 (21)
23. de gode		
bondesamfunnet	14. en varme leilighet	2. det fritt aviser
	18. et nasjonal symbol -	
27. et tøff problem	et	3. den mørke vinter
28. de greie kirken	19. et blå bygg -et	7. en aktivt forfatter
	23. den god	
35. de fine turister	bondesamfunnet	17. det nytt språket
38. de tunge bøker	24. sentral saken	18. en nasjonale symbol
39. et vill fjell	25. spesielt demokratiet	19. det blått bygg
40. det åpne		23. den gode
oppførselen	27. et tøff problem	bondesamfunnet
42. annet tema	31. kaldt været	24. den sentral saken
44. de vanskelige	32. den morsome	
oppgaver	programmene	26. friske nyhetene
	34. et effektivt	
45. de grønne laget	oppdragelse -en	28. en greie kirke
		32. den morsomme
48. en svak intrykk	36. det sterkt diskusjonen	progrmene
51. et gammel våpen	37. et rikt kultur -en	34. en effektivt oppdragelse
52. de lave svaret	43. stor forfatteren	40. den åpen oppførsel
	44. det vanskelige	
53. et ivrig turgåere	oppgaver	45. det grønt laget
56. en enorm ønske	45. den grønne laget	46. en formelle festen
57. et liten artikkel	50. en aktuelt artikkel -en	48. et svak intrykk
58. de normale klimaet	52. det lav svaret	49. den dypt forskjellige
	55. en offisielle	
59. de høye studenter	undersøkelse	50. en aktuelle artikklen
	57. den liten artikkelen	57. den liten artikklen
	58. det normalt klimaet	58. det normalt klimaet
		60. de positive ordet

6. Instructions for the on-line DP production task

Rettledning: Du vil få høre noen substantivfraser på norsk, der adjektivet *gammel* står til substantivet. Enkelte av frasene er velformede, mens andre ikke er det. Det eneste du skal gjøre, er å repetere substantivet som står etter adjektivet, så nøyaktig og så raskt som du kan. Hele oppgaven tar omtrent 9 minutter. Det vil bli en kort pause hvert tredje minutt. Her er tre eksempler du kan øve deg på:

Første eksempel: *det gamle navnet*

Andre eksempel: *den gamle huset*

Tredje eksempel: *de gamle bilen*

7. Test items, noun frequency, noun duration in the stimuli of the auditory naming task (gender-marking)

NO.	Stimuli	Noun frequency	Noun duration
1	den gamle bordet	1208	482
2	den gamle landet	9541	601
3	den gamle folket	7055	669
4	det gamle krisen	762	620
5	den gamle arbeidet	2150	670
6	det gamle kampen	6272	717
7	den gamle navnet	3001	624
8	det gamle vennen	1909	543
9	det gamle planen	3010	547
10	den gamle valget	2966	587
11	den gamle medlemmet	2572	752
12	det gamle fordelen	820	750
13	det gamle brannen	2834	543
14	det gamle bilen	4136	511
15	den gamle svaret	2160	646
16	det gamle legen	1835	635
17	det gamle avtalen	1901	782
18	den gamle banken	2493	633
19	det gamle produktet	904	721
20	det gamle huset	3826	610
21	det gamle brevet	1886	538
22	det gamle firmaet	850	722
23	det gamle salget	2314	670
24	det gamle rommet	2226	519
25	den gamle gutten	2308	515
26	det gamle livet	7064	617
27	den gamle byen	5690	502
28	den gamle saken	9877	676
29	den gamle ballen	1186	551
30	det gamle kurset	2164	608
31	den gamle løsningen	2322	818
32	det gamle tilfellet	1721	671
33	den gamle meningen	2003	704
34	den gamle veien	4681	582

35	den gamle gaven	787	571
----	-----------------	-----	-----

8. Test items, noun frequency, noun duration in the stimuli of the auditory naming task (number-marking)

NO.	Stimuli	Noun frequency	Noun duration
1	de gamle kommunen	6081	580
2	de gamle systemet	1209	774
3	de gamle regjeringen	3936	751
4	de gamle skaden	1507	660
5	de gamle myndigheten	3823	847
6	de gamle feilen	1896	583
7	de gamle skolen	5363	638
8	de gamle tallet	1856	542
9	de gamle våpenet	1210	632
10	de gamle problemet	5076	702
11	de gamle kongen	1471	514
12	de gamle forslaget	3711	760
13	de gamle tilbudet	2671	708
14	de gamle leiligheten	1420	887
15	de gamle debatten	1479	617
16	de gamle sykehuset	3154	894
17	de gamle lånet	1585	547
18	de gamle spørsmålet	4583	768
19	det gamle tilfellet	1721	671
20	det gamle produktet	904	721
21	det gamle huset	3826	610
22	det gamle brevet	1886	538
23	det gamle firmaet	850	722
24	det gamle salget	2314	670
25	det gamle rommet	2226	519
26	det gamle kurset	2164	608
27	det gamle livet	7064	617
28	den gamle ballen	1186	551
29	den gamle byen	5690	502
30	den gamle løsningen	2322	818
31	den gamle saken	9877	676

32	den gamle meningen	2003	704
33	den gamle veien	4681	582
34	den gamle gaven	787	571
35	den gamle banken	2493	633
36	den gamle gutten	2308	515

9. Test items, noun frequency, noun duration in the stimuli of the auditory naming task (definiteness-marking)

NO.	Stimuli	Noun frequency	Noun duration
1	de gamle situasjoner	2725	871
2	de gamle planer	3010	498
3	de gamle oppgaver	2714	718
4	de gamle familier	3156	767
5	de gamle kontorer	1225	746
6	de gamle muligheter	5081	801
7	de gamle linjer	1819	693
8	de gamle konserter	1285	760
9	de gamle ulykker	1482	658
10	de gamle historier	3440	727
11	de gamle tekster	4803	605
12	de gamle resultater	3434	906
13	de gamle former	4125	623
14	de gamle pasienter	1380	816
15	de gamle temaer	1269	631
16	de gamle jenter	4659	695
17	de gamle verdier	1293	678
18	de gamle kirker	5262	723
19	de gamle årsakene	2201	796
20	de gamle avisene	3234	714
21	de gamle uttrykkene	1671	762
22	de gamle grensene	2178	673
23	de gamle krigene	1813	648
24	de gamle helgene	3160	677
25	de gamle elevene	2545	679
26	de gamle skipene	1613	673
27	de gamle dyrene	1576	548

28	de gamle delene	6940	626
29	de gamle gruppene	3625	665
30	de gamle prisene	5495	666
31	de gamle klassene	2399	662
32	de gamle varene	2736	679
33	de gamle rekkene	3597	697
34	de gamle dørene	1805	641
35	de gamle gatene	2329	633
36	de gamle spillene	1706	705