Endstate L2 Competence: Does Perception Parallel Production?

Fufen Jin, Tor A. Åfarli, and Wim A. van Dommelen
Norwegian University of Science and Technology (NTNU)

1. Introduction

A fundamental concern of second language (L2) research is whether or not post-critical L2 learners can ultimately achieve linguistic competence that is identical to that of a native speaker. "Linguistic competence" in our view entails not only grammatical knowledge but also grammatical processing. In this paper we wish to explore the interrelation between the two aspects of linguistic competence in endstate L2 learners, an issue that has not received much attention in previous L2 research. Specifically we ask the question whether L2 learners' parsing abilities are directly related to their grammatical knowledge of the target language. In other words, does perception parallel production? A related question is the role of L1 in both L2 production and perception. Does L1 transfer play a role in the endstate L2 interlanguage grammar and in L2 processing? This paper attempts to address these issues through examining the production and perception of Norwegian DP internal agreement by endstate L2 speakers of Norwegian, who are from three typologically different L1 backgrounds: Italian/Spanish, English, and Chinese. In the following section we will present some crosslinguistic facts and theoretical assumptions regarding the DP structures of the target and the source languages.

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

2.1 DP internal agreement in Norwegian

Because the focus of our study is the production and perception of Norwegian DP internal agreement, it is important to give a descriptive overview of how DP internal agreement is realized in the target language. The Norwegian language is characterized with a rather complex inflectional system in its nominal domain, which encodes gender, number and definiteness. In addition, there are DP internal agreement between determiners, adjectives and nouns in terms of gender, number, and quite uniquely, definiteness. The Norwegian DP internal agreement is illustrated in (1) and (2) below:[1]

(1) a. et hus
D-neut.sg. indef. house-neut.sg. indef.
'a house'

b. hus-et
house-neut.sg.def.
'the house'

(2) a. et gammel-t hus
D-neut.sg. indef. old-neut.sg.indef house-neut.sg. indef.
'an old house'

b. det gaml-e hus-et
D-neut.sg.def. old-*w* house-neut.sg.def.
'the old house'

[1] The following symbols are used for the annotation throughout the paper. *D* stands for prenominal determiners; *neut* for neuter gender, *com* for common gender, *sg* for singular number, *pl* for plural number; *def* for definite, *indef* for indefinite; *w* for the weak inflection on adjectives.

 Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA), ed. Alyona Belikova et al., 200-211. Somerville, MA: Cascadilla Proceedings Project.

As we can see, in simple Norwegian DPs the indefinite determiner is realized as a prenominal free morpheme (cf. 1a), whereas the definite determiner is realized as a suffixed bound morpheme (cf. 1b). When a DP contains an attributive adjective, the adjective always appears in a prenominal position. For indefinite DPs, the insertion of an adjective has no effect on the determiner (cf. 2*a*). For definite DPs, however, the inclusion of an adjective results in the so-called double definiteness construction (cf. 2*b*), where there is a co-existence of a suffixed determiner (*-et*) on the noun, and a free prenominal determiner (*det*), and the two determiners agree with each other in gender and number. The attributive adjective in the double definiteness construction appears with an invariant (weak) inflection *–e*. Table 1 lists the regular inflectional paradigm of modified Norwegian DPs with all possible feature combinations. Note that Norwegian distinguishes neuter gender from non-neuter (or common) gender, and that the gender information on the determiner is neutralized in plural form. [2]

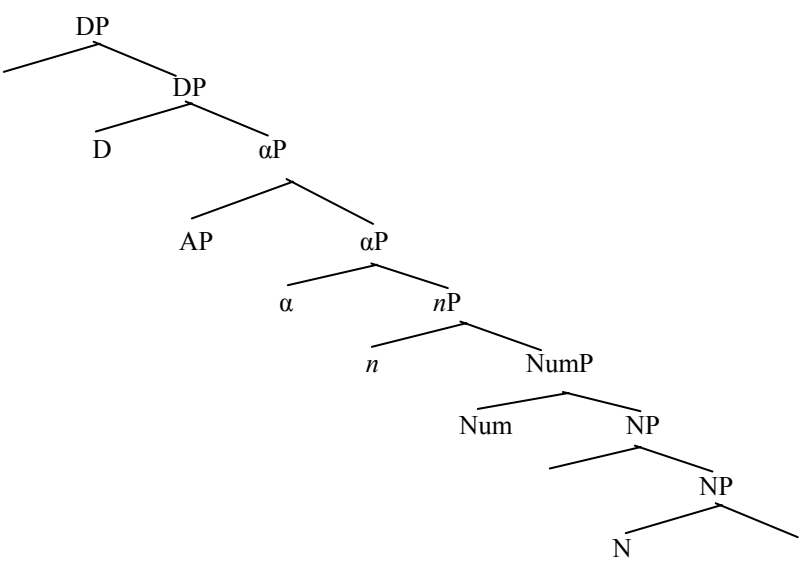
Table 1. Inflectional paradigm of modified Norwegian DPs

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

2.2 Theoretical assumptions

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 (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, so the analysis of Norwegian DPs in our study is based on her proposal. Following Julien, we assume that the (extended) Norwegian DPs have the syntactic structure shown in (3).

(3)



As we can see from (3), 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. Just as V obligatorily moves to *v*, N obligatorily moves to

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

n, via the Num head. Moreover, *n* hosts the suffixed definite determiner in Norwegian. When an adjective is included, *n*P merges with an *α* head, which projects *α*P, with the adjectival phrase (AP) sitting in its Spec. Finally, on the topmost layer, we find DP, with D hosting the free nominal determiners.

Under the assumption that syntactic features are morphologically expressed, 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 and *α*. 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 uninterpretable features on D and *α* heads. The parameterized uninterpretable features in target and source languages are summarized in table 2, based on the facts whether or not the languages have overt gender, number and definiteness agreement between determiners, adjectives, and nouns. [3]

Table 2. Parameterized uninterpretable features in target and source languages

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

Key: + present in language; - absent in language.

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 uninterpretable features either on the D or on the *α* head. Given these parametric differences between the target and the source languages, we are in a position to examine whether there are L1-L2 differences, and differences among the three L2 groups in production and perception of the Norwegian DP internal agreement.

3. Method

3.1 Participants

Six English informants, five Chinese informants, and five Italian/Spanish (2 Italian; 3 Spanish) informants were selected according to the following criteria:

i) They had to have residence in Norway for at least eight years;
ii) Their self-rated oral comprehension and oral production of Norwegian were both over 8, on a 10 point scale;
iii) They acquired Norwegian as an adult (>13 years old), and had not learn any other Scandinavian languages before age 13;
iv) For English-speaking informants, they did not acquire any gender-marking languages as a child (ages from 0 to13);
v) For Chinese speaking informants, they did not acquire any gender or number marking languages as a child (ages from 0 to13).

All the L2 informants were interviewed concerning their biographical information and language learning background. In addition, their knowledge of the Norwegian language was tested, using a Norwegian Proficiency Test (NPT, which was a Close test with 50 blanks). Details of the informants' information and their NPT scores are given in table 3. All of them had lived, worked or studied in a

[3] We use [*u*] to indicate that a given feature is unvalued.

fundamentally Norwegian-speaking environment. They all had great motivation to acquire Norwegian, and appeared to be very successful, as can be seen from their high scores of Norwegian proficiency test (≥ 46 out of 50) and their reported highly proficient Norwegian oral comprehension and production (≥ 8 on a 10 point scale).

Table 3. L2 informants' background information and language proficiency

ID	M/F	AGE	LOR	ASSN	NOC	NOP	NPT
E1	F	60	32	21	9.5	9.5	50
E2	M	34	12	21	8.5	8	47
E3	M	36	11	25	8	8	48
E4	F	44	19	15	10	10	50
E5	M	37	10	27	8	8.5	47
E6	M	32	8	23	9.5	9.5	49
R1	M	44	11	29	9	9	48
R2	F	40	12	24	9.5	10	47
R3	M	56	20	32	9	9	48
R4	F	53	25	28	10	10	47
R5	F	30	8	22	9	8.5	48
C1	F	31	17	14	9.5	9.5	48
C2	M	25	9	16	8	9	48
C3	M	44	12	32	8	8	46
C4	F	37	10	27	8	8	47
C5	F	35	10	25	8	9	47
Mean (sd.)		39.8 (9.5)	14.1 (6.6)	23.8 (5.3)	8.84 (.74)	8.97 (.72)	47.8 (1.1)

Key: M/F=sex; LOR= length of residence in Norway; AGE=age at the time of testing; ASSN=age started speaking Norwegian; AOB= age of onset of L1/Norwegian bilingualism; NOC=self rated Norwegian oral comprehension; NOP= self rated Norwegian oral production, NPT=Norwegian proficiency test.

In addition, a control group of fourteen native speakers of Norwegian (6 males, 8 females) was recruited. The L1 participants fell within the age range of their L2 counterparts. Their mean age was 36.4 (sd. 10.6), ranging from 25 to 64. No informants had any hearing impairments.

3.2 Experimental tasks

The informants were tested on two tasks: an online DP production task and an auditory naming task. The purpose of the production task was to determine whether the L2 participants can have native-like performance regarding the DP internal agreement. In particular, will the [*u*NUM], [*u*GEN], [*u*DEF] features be present in the interlanguage grammar regardless of presence or absence of the corresponding features in the L1s? The on-line mode was selected in order to reduce the degree to which the participants have access to their metalinguistic knowledge. The purpose of the auditory naming task was to examine whether the L2 speakers' sensitivity (or insensitivity) to the agreement cues is directly related to how well they produce the correct agreements. In other words, does production parallel perception? By the auditory naming task, we can examine the effects of the prime on the noun targets that are syntactically related to the prime. In this task we used D as a prime that is either concordant or discordant with noun targets in terms of gender, number and definiteness, respectively. If significant concordance-discordance difference in reaction time is found, this will be an indication that the participant is sensitive to the cues on the prime.

3.2.1 Materials and procedures

3.2.1.1 Experiment 1: On-line modified DP elicitation task

Materials: The experiment was made up of 60 short sentences (average length=5.2), all containing a simple nominal phrase. According to the gender, number and definiteness specifications of the nouns,

the sentences were broken into the following 6 conditions, 10 in each condition. This is illustrated in table 4.

Table 4. Types of DPs and illustrations in experiment one (with target DPs in italicized form)

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

Procedure: The informants first heard a short sentence presented to them over headphones. Simultaneously they saw on the computer screen a noun (in its bare form), which had occurred in the sentence, followed by an adjective (put in brackets). 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 a sentence in table 4 as an example, the informants hear '*Klimaet er ganske bra*' and see on the computer screen '*klima* (normal)'. The expected response is '*Det normale klimaet er ganske bra'*. Informants were tested individually in a phonetic laboratory. The test items were mixed and randomized, with an interstimulus interval set at 5 seconds. The order of presenting the stimuli was the same for all informants.

3.2.1.2 Experiment 2: auditory naming task

Materials and Design: Stimuli for the experiment were auditory Norwegian DPs of the "double definiteness" construction, which were made up of a prenominal determiner, an adjective (a single adjective *gamle* 'old' was used in this experiment), and a noun with a suffixed determiner. Depending on the grammatical agreement features, the test items were divided into 3 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 head noun. By discordant, we mean that the determiner has a mismatch of a single feature (gender, number, or definiteness) with the head noun. Examples of concordant and discordant test items in terms of gender, number and definiteness are illustrated in table 5 below (with mismatched features in bold form).

Table 5. Sample stimuli

	CONCORDANT	DISCORDANT
GEN	den gamle bil-en D-com.sg.def. car-com.sg.def.	den gaml-e bord-et D-**com**.sg.def. table-**neut**.sg.def.
NUM	det gaml-e kurs-et D-neut.sg.def. course-neut.sg.def.	de gaml-e hus-et D-**pl**.def. house-neut.**sg**.def.
DEF	de gaml-e krig-ene D-pl.def. war-pl.def.	de gaml-e plan-er D-pl.**def.** plan-pl.**indef.**

All in all, 90 different nouns were selected for this study. There was no repetition of nouns in the test items. We also tried to counterbalance the word frequency (based on *the Oslo Corpus of Tagged Norwegian Text* http://www.hf.uio.no/tekstlab/frekvensordlister/index.html) and the syllable length of nouns in discordant and concordant conditions. This was achieved in gender and definiteness marking

groups, but in number-marking, only word frequency was counterbalanced, at the cost of syllable length. All the test items were audio-recorded by a female native speaker of Norwegian 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 wished to ensure that the noun targets are preceded by determiner-adjective pairs of a similar duration. Also care was taken to ensure natural transitions between the words and to achieve appropriate amplitude relations. The test items were mixed and randomized, with an interstimulus interval set at 3.5 seconds. The order of presenting the stimuli was the same for all informants.

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 were 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. It was emphasized that the noun targets should be repeated in the same form as they appear in the recording; no correction should be made in repeating the words. 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 a time pressure. Norwegian was used throughout the testing session. The test took about 9 minutes, and there was a short break every 3 minutes. 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.

4. Results

4.1 Experiment 1

We calculated the percentage of target-like DPs out of all the modified DPs produced by each of the individual participants. As expected, no agreement errors were made by the native informants. The accuracy rate of the DPs produced by the individual L2 participant is given in figure 1.

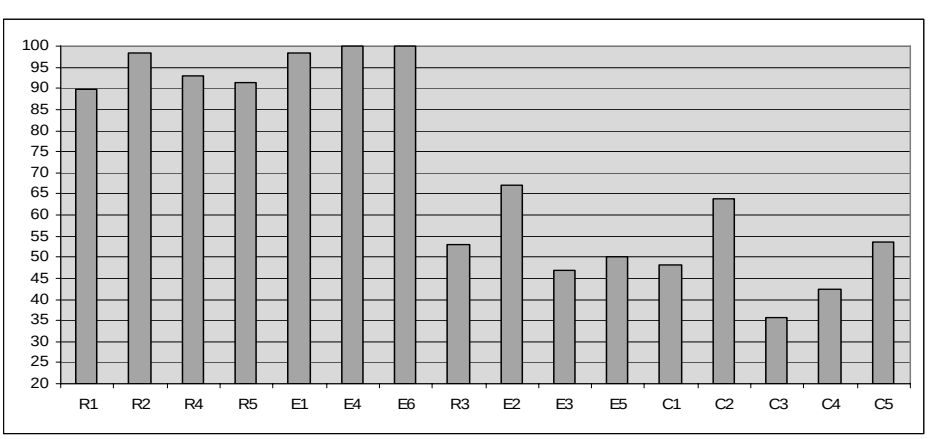


Figure 1. Accuracy rate of the DPs produced by the individual informants

It seems that the task divided the L2 participants into two groups. One group (including 4 informants from the Romance group and 3 informants from the English group) had native-like

performance, with correct percentage of 89% or above; while the remaining nine informants (including all informants from the Chinese group, 3 informants from the English group, and one informant from the Romance group) produced correct DP internal agreement more or less at a chance level. The Chi-squared test confirmed our observation. Significant difference was found between R1, who has the lowest correct percentage among the native-like performers, and E2, who had the highest percentage among the non-native-like performers ($X^2 = 9.624$, df =1, p=.002).

Now we rearrange the L2 participants according to their performance in experiment 1. Those who performed native-like were assigned to group 1 (G1), and those who performed at the chance level to group 2(G2). We will look at the agreement errors made by G2 participants and find out whether there is an L1 effect at work. We divided the errors 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 (%) in each of the categories was calculated for all the three L2 groups within G2, and the results were presented in figure 2.

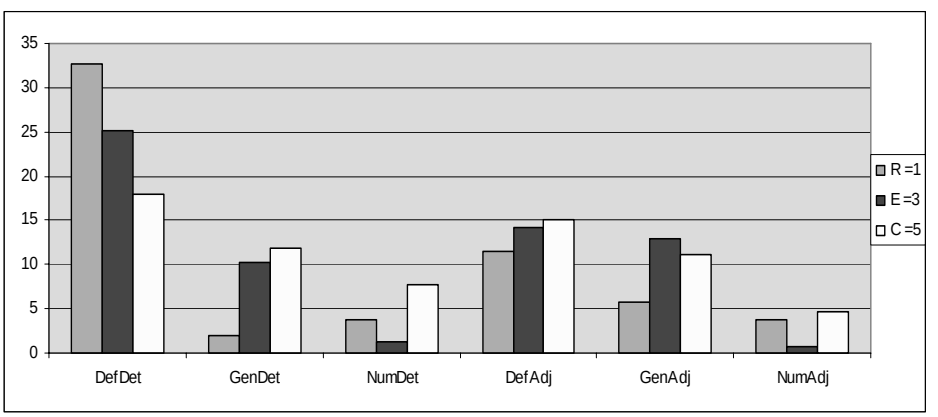


Figure 2. Error rate (%) in each of the categories for the three L2 groups within G2

All the three L2 groups made most errors in definiteness agreement both on determiners (32.7%, 25.2%, 17.9% for the Romance, the English and the Chinese group, respectively) and on adjectives (11.5%, 14.2%, 15.1% for the Romance, the English and the Chinese group, respectively). The Romance group had the smallest gender agreement error rate among the three groups. The Chinese group made more number agreement errors than the English and the Romance groups (7.7% as against 1.7% and 3.8%). It is clear from these results that those who made agreement errors did show an L1 effect—those uninterpretable features which are not instantiated in L1s (either root D or root α) were areas of persistent difficulty in these L2 speakers.

4.2 Experiment 2

4.2.1 L1 participants

Since there were seven L2 participants in G1 who performed native-like in the on-line DP production task, and since the auditory naming task aimed to examine whether these L2 participants use agreement cues in perception the way natives do, we selected from the control group seven native speakers, who were matched on gender and age with the G1 (mean ages were 43.2 for G1 and 41.6 for the control group), in order to render the L1 and L2 groups comparable. Mean RTs of the L1 participants were submitted to analysis of variance (ANOVA). Table 6 shows the mean RTs in each cell.[4]

[4] As syllable length between the concordant and the discordant conditions was different in the number-marking group, we selected 11 items from each condition, with their word frequency and syllable length counterbalanced. All data analysis regarding number-marking used the counterbalanced data.

Table 6. L1 participants' mean RTs to gender, number, and definiteness markings in each condition (RT in milliseconds; standard deviation in brackets)

	Grammatical agreement features			
Determiners	Gender	Number	Definiteness	Total
Concordant	569	582	611	588
	(53)	(61)	(48)	(55)
Discordant	647	644	682	659
	(80)	(70)	(74)	(76)
Difference	78	62	71	71

A 2(concordant vs. discordant) × 3 (gender, number, and definiteness) ANOVA was conducted. The ANOVA yielded a main effect of concordance, [F (1, 86) =25.47, p<.001]. Collapsed over all three agreement features, targets primed by a concordant D were responded to about 71ms faster than those primed by a discordant D. The interaction between the two factors (cue type vs. agreement features) was not significant [F (2, 86) = .097, p=.907], indicating that the processing pattern was the same for all the three grammatical agreement features.

Separate analyses were carried out for the three agreement feature groups. The 78ms concordance-discordance difference in gender-marking, 62ms difference in number-marking, and 71ms difference in definiteness-marking were all significant [$t_{gen}(33)=3.41$, p=.002; $t_{num}(20)=2.22$, p=.038; $t_{def}(34)=2.78$, p=.009]. 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, which demonstrates that native speakers are sensitive to gender, number and definiteness cues on the determiners when processing Norwegian.

4.2.2 L2 participants

The same data analysis procedure was applied first to G1 L2 participants and then to G2 L2 participants. Table 7 and 8 give the mean RTs in each cell for G1 and G2 respectively.

Table 7. G1 L2 participants' mean RTs to gender, number, and definiteness markings in each condition (RT in milliseconds; standard deviation in brackets)

	Grammatical agreement features			
Determiners	Gender	Number	Definiteness	Total
Concordant	565	569	606	582
	(58)	(66)	(55)	(61)
Discordant	572	611	629	603
	(75)	(60)	(74)	(73)
Difference	7	42	23	21

Table 8. G2 L2 participants' mean RTs to gender, number, and definiteness markings in each condition (RT in milliseconds; standard deviation in brackets)

	Grammatical agreement features			
Determiners	Gender	Number	Definiteness	Total
Concordant	576	629	598	597
	(67)	(74)	(59)	(67)
Discordant	583	642	632	616
	(66)	(67)	(75)	(74)
Difference	7	13	34	26

Two striking facts can be observed from the data above. First, the overall priming effects were quite comparable between the two L2 groups. It seems that native-like production of Norwegian DP internal agreement is not closely 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 were 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 [all ps >.05]. These results thus clearly indicate that L2 participants were not sensitive to any of the agreement cues on D, whether or not they performed native-like in the production task.

5. Discussion

5.1 An account for L1-L2 difference in processing mechanism

Results from the auditory naming task suggest that adult L2 learners do not completely acquire the same processing mechanism as the natives, despite their native-like performance in the production task. Similar findings were reported in Guillelmon and Grosjean (2001), who used an auditory naming task to examine how late English-French bilinguals react to gender marking when processing French. They found that late bilinguals were totally insensitive to gender marking in perception, in contrast to native French controls, who showed clear facilitation and inhibition effects. These findings together lend support to the *shallow structure hypothesis* (SSH) proposed by Clahsen and Felser (2006), who claim that the representations adult L2 learners compute during processing contain less syntactic detail than those of child and adult native speakers. Here we are concerned with the 'syntactic detail' that is available for the native speakers, but inaccessible for adult L2 learners in processing Norwegian DP internal agreement. If we consider the on-line nature of the task, which requires the informants to respond as quickly as possible, the interesting question to ask is: "why couldn't the native speakers simply ignore the agreement cues on the primes altogether?", as they could observe that the primes offered completely unreliable information about the identity of noun targets. 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)?

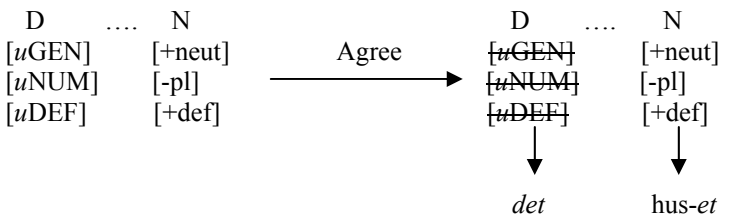
To answer this question, we propose a model of agreement processing, which incorporates a Probe-Goal account of grammatical agreement developed in recent work of the Minimalist Program (MP) (Chomsky, 1995, 1999, 2000). It is assumed that for native speakers, syntactic computation takes place during the syntactic formulation of an utterance and during its parsing as well (Correa et al. 2005; Fong 2004; Phillips 2004). In light of the MP, uninterpretable features, which enter the derivation without values, must be valued and eliminated for the purpose of convergence. This is done via an operation called Agree, which establishes a Probe-Goal relation between the uninterpretable features and the matching interpretable features. Applying this theory to Norwegian, we can depict the structure building process of the DP *det gamle huset* in the following way. [5] First the correct components for assembly are laid out. In this case, the components include D with uninterpretable gender ([*u*GEN]), number ([*u*NUM]), and definiteness features ([*u*DEF]), the 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. The uninterpretable features in 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 feature values on D spell out as *det*.

This process can be schematically represented as follows:

[5] For the purpose of this paper, we only focus on the agreement between D and N. There may well be a Probe-Goal relation between α (with an Adjective Phrase in its Spec) and N as well (see Julien 2005 for a detailed account), but this operation presumably has no effect on processing in our case, as we used an invariant adjective (i.e. *gamle*) in all our test items.

(4) 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 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 (e.g. *den, det* or *de*) is discovered. Once a DP is analyzed as such, the uninterpretable features on D will drive the parsing process by establishing a Probe-Goal relation between D and N through the matching of features, and the operation Agree applies automatically. 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. The Probe-Goal relation between D and N explains why the information on D affects the processing of N.

To account for the results obtained from the L2 participants, we suggest, in line with the shallow structure hypothesis, L2 participants process Norwegian DPs without invoking full syntactic computation. In this case, it is the operation Agree that does not apply in L2 processing. Since the Probe-Goal relation between D and N is not established, the information on D presumably has no effects on the processing of N. Thus the L2 participants were overall insensitive to the agreement cues on D.

5.2 L1 transfer in DP production and perception

The parametric differences between the target and the source languages regarding uninterpretable features enable us to examine L1 transfer effects in L2 production and processing. In the field of theoretical second language acquisition, there has been considerable debate on the availability of the parameterized uninterpretable features to adult endstate L2 speakers. Researchers assuming the *full transfer full access hypothesis* (e.g. Schwartz & Sprouse 1996; White 2003; White, et al. 2004) argue that all features are fully available to the adult L2 learners, irrespective of their status in the L1s. Researchers assuming the *failed functional features hypothesis* (e.g. Hawkins 2004; Hawkins & Franceschina 2003), on the other hand, contend that those uninterpretable syntactic features which are not instantiated in the L1s disappear beyond a critical period. Hence, they are no longer available to L2 speakers.

Our findings from the DP production task were not fully compatible with either of the models. It appears only the Chinese informants behaved in the manner predicted by the failed functional features hypothesis. Neither the Romance nor the English group, however, behaved uniformly. On the one hand we found that four Italian/Spanish informants and three English informants performed native-like, irrespective of the absence of uninterpretable definiteness feature in their L1s (and for English informants, [*u*GEN] feature is absent in their L1 as well), on the other hand, there were one Spanish informant and three English informants who produced DP internal agreement with an accuracy rate more or less at a chance level. More fine-grained analysis of agreement errors made by these informants did show an L1 effect— those uninterpretable features which are not instantiated in L1s (either root D or root α) were areas of persistent difficulty for these L2 speakers. Our findings suggest that there are individual differences as to the availability of the parameterized uninterpretable features to endstate L2 speakers. Some can acquire the parametric uninterpretable features and have native-like performance in

terms of DP internal agreement, whereas others cannot acquire the uninterpretable features that are not instantiated in their L1s. But one thing is certain, and quite compatible with the failed functional features hypothesis. This is, if the L2 speakers do 'fail', it is the parameterized uninterpretable features that are subject to a critical period.

L1 transfer in L2 processing is also an issue of great controversy (see Clahsen and Felser 2006 for a review). Based on SSH, Clahsen and Felser (2006) claim that L1 transfer influences L2 processing only indirectly. This 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 there is conflicting evidence as to the role of L1 transfer in L2 processing within local grammatical domains. For example, in an ERP study, Sabourin (2003) found that German-speaking learners of Dutch were sensitive to gender violations in Dutch, while Romance- and English-speaking learners of Dutch were not, yielding evidence of L1 transfer (among the source languages, only German has a gender system that is congruent to Dutch gender system). Tokowicz and MacWhinney (2005), on the other hand, observed that English learners of L2 Spanish were implicitly sensitive to determiner gender violations but not to determiner number violations.

Due to the small sample size of the G1 participants whose L1s have gender and gender agreement, we were not in the position to examine the L1 transfer effect of gender in L2 processing. But it is possible for us to look at the transfer effects of number and definiteness features, as Italian/Spanish resembles English in having overt number agreement and no definiteness agreement between determiners and nouns. We have learned that G1 participants' naming time difference between the concordant and discordant conditions were not significant either in number marking or in definiteness marking, indicating that these L2 speakers were insensitive to either number or definiteness cues on D. It is evident from these results that L1 transfer plays no role in the processing of L2 Norwegian. We can conclude that even those endstate L2 speakers who had acquired native-like grammatical knowledge with respect to DP internal agreement, did not completely acquire the same processing mechanism as the natives. The findings provide evidence that SSH can be extended to local domains, not being restricted to non-local domains. But whether or not this phenomenon is task specific (i.e. only limited to auditory naming task) has to be left for future research.

6. Conclusion

The primary goal of this research is to determine the interrelation between production and perception. Having tested L2 speakers' production and perception of Norwegian DP internal agreement, we found that L2 speakers could not use agreement cues on D in perception, irrespective of their ability to produce Norwegian DP internal agreement in a native-like manner. Our findings are in line with shallow structure processing hypothesis (SSH) proposed by Clahsen and Felser (2006), who argue for L1/L2 differences in processing. But contrary to their claim that the SSH is restricted to non-local domains, the results from our studies suggest that it can apply to L2 processing within local domains as well. Our findings also provide evidence that even highly proficient, endstate L2 speakers are subject to the SSH in processing an L2. A cautious note to be taken is that so far we are 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.

References

Bates, Elizabeth, Devescovi, A., Hernandez, A and Pizzamiglio, L. (1996). Gender priming in Italian. *Perception and Psychophysics*, 58, 992-1004.

Boersma, Paul and Weenink, D. (2006). Praat: doing phonetics by computer (Version 4.4.11) [Computer program]. Retrieved February 23, 2006, from http://www.praat.org/.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam (1999). Derivation by phase. *MIT Occasional Papers in Linguistics 18*. Cambridge, Mass.: MITWPL.

Chomsky, Noam (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* (pp. 89-155). Cambridge, Mass.: The MIT Press.

Clahsen, Harald & Felser C. (2006). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, 27, 107-126.

Correa, Leticia-S., Augusto M. & J. Ferrari-Neto. (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 2005.

Delsing, Lars-Olof (1993). The internal structure of noun phrases in the Scandinavian languages. Unpublished doctoral dissertation, Lund University.

Fong, Sandiway (2004). Computation with probes and goals: A parsing perspective. In A.M. Di Sciullo & R. Delmonte (eds.), *UG and External Systems* (pp. 311-334). Amsterdam: John Benjamins.

Guillelmon, Delphine & Grosjean, F. (2001). The gender marking effect in spoken word recognition: the case of bilinguals. *Memory & Cognition*, 29,503-511.

Hawkins, Roger (2004). Revisiting Wh-movement: the availability of an uninterpretable [wh] feature in interlanguage grammars. *Cascadilla Proceedings Project.* Somerville, MA: Cascadilla Press.

Hawkins, Roger & Franceschina, F. (2003). Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. In Paradis, J. & Prévost, P. (eds.), *The acquisition of French in different contexts* (pp.175-205). Amsterdam: John Benjamins.

Julien, Marit (2003). Double definiteness in Scandinavian. *Nordlyd*, 31, 230-244.

Julien, Marit (2005). *Nominal Phrases from a Scandinavian Perspective*. Amsterdam: Benjamins.

Kester, Ellen-Petra (1996). The nature of adjectival inflection. Unpublished doctoral dissertation, Utrecht University.

Marinis, Theodore, Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, 27, 53-78.

Phillips, Colin (2004). Linguistics and linking problems. In M. Rice & S. Warren (eds.), *Developmental Language Disorders: from Phenotypes to Etiologies* (pp.241-290). Mahwah, N.J.: Lawrence Erlbaum.

Sabourin, Laura (2003). Grammatical gender and second language processing: an EPR study. Unpublished doctoral dissertation, University of Groningen.

Schwartz, Bonnie & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access hypothesis. *Second Language Research,* 12, 40-72.

Tokowicz, Natasha, & 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.

Vangsnes, Øystein A. (1999). The identification of functional architecture. Unpublished doctoral dissertation, University of Bergen.

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

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

Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)

edited by Alyona Belikova,
Luisa Meroni, and Mari Umeda

Cascadilla Proceedings Project Somerville, MA 2007

Copyright information

ISBN 978-1-57473-419-5 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Jin, Fufen, Tor A. Åfarli, and Wim A. van Dommelen. 2007. Endstate L2 Competence: Does Perception Parallel Production? In *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, ed. Alyona Belikova, Luisa Meroni, and Mari Umeda, 200-211. Somerville, MA: Cascadilla Proceedings Project.

or:

Jin, Fufen, Tor A. Åfarli, and Wim A. van Dommelen. 2007. Endstate L2 Competence: Does Perception Parallel Production? In *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, ed. Alyona Belikova, Luisa Meroni, and Mari Umeda, 200-211. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1561.