Sensitivity to DP-Internal Agreement Violations in L2 Grammar

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

1. Introduction

The paper aims to explore a fundamental question concerning whether or not endstate L2 speakers can attain nativelike processing capacity, an issue that has not received much attention in the previous research on L2 ultimate attainment. The domain to investigate L2 processing capacity is agreement dependencies in Norwegian determiner phrases (Determiner-Adjective-Noun). Norwegian has 'double definiteness' construction, where there is visible agreement between the prenominal determiner (D) and the suffixed determiner on the noun (N), as well as gender and number agreement (see section 2 for details). L2 participants were speakers of an L1 that displays gender and number agreement (Italian/Spanish), only number agreement (English), and no agreement (Chinese).

The experimental technique used in order to investigate L2 speakers' processing abilities is called auditory naming (alternatively called cued shadowing, or spoken word repetition, Bates & Liu 1997), in which subjects are asked to listen to pairs of phrases and repeat a target word embedded in the carrier phrase. This technique has been used with great success to examine the degree to which L1 speakers are sensitive to local domain syntactic and semantic violations (see Bates and Liu, for a review).

This paper proceeds as follows. Section 2 introduces syntactic properties of Norwegian 'double definiteness' construction, and sum up the parametric differences between the target language and the source languages. Section 3 offers a brief review of the literature on L1 and L2 processing of local domain feature mismatches. 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. This model serves as the starting point for designing the auditory naming experiment, the details of which are presented in section 4. Subsequently, in section 5, we present results from the experiments, first for L1 participants, and then for L2 participants. In section 6, the observed L1 vs. L2 processing differences are discussed in relation to processing theories, in particular Shallow Structure Hypothesis (Clahsen and Felser 2006). Finally, a conclusion will be reached in section 7.

2. Double definiteness construction in Norwegian

In Norwegian, the definite determiner in an unmodified DP is realized as a suffixed bound morpheme, which also encodes gender (neuter vs. common gender) and number information (cf. 1a). [1,2] However, when an attributive adjective is inserted in the definite DPs, it requires the co-occurrence of both a suffixed determiner and a free prenominal determiner; the two definite determiners agree in terms of gender and number as well (cf. 1b). 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 'double definiteness' construction. The attributive adjective in the 'double definiteness' construction appears with an invariant (weak) inflection *–e*.

[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.
[2] There are two official forms of written Norwegian: *Bokmål* 'book language' and *Nynorsk* 'new Norwegian'. The written form used throughout this paper is of the *Bokmål* variety.

 Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008), ed. Jean Crawford et al., 116-126. Somerville, MA: Cascadilla Proceedings Project.

(1) a. hus-et
house-neut.sg.def.
'the house'

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

Julien (2003, 2005) provides a thorough investigation of Scandinavian DPs within the minimalist framework (Chomsky 1995). Relevant to our current study is her proposal that the D head in Norwegian contains uninterpretable gender, number, and definiteness features. These features enter into derivation unvalued. We thus refer to them as [*u*GEN], [*u*NUM], and [*u*DEF], respectively, using [*u*] to represent unvalued features. These unvalued features must be valued and deleted by establishing an agreement relationship with the matching interpretable features down in the N head. Julien also 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. Following this universalist claim, we assume that the DP structures of Italian/Spanish, English, and Chinese are fundamentally the same with that of Norwegian. Variations among these languages are attributed to presence or absence of uninterpretable features on D. The parameterized uninterpretable features in target and source languages are summarized in table 1, based on the facts whether or not the languages have overt gender, number, and definiteness agreements between determiners and nouns.

Table 1. The parametrized uninterpretable features on D

	D		
	[*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 on the D head; Italian/Spanish has [*u*NUM] and [*u*GEN], but not [*u*DEF] feature; the only uninterpretable feature English has is a [*u*NUM] feature; Chinese has none of the uninterpretable features. 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 processing the Norwegian DP internal agreement.

3. 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 (see for example, Grosjean et al. 1994, for French; Bates et al. 1996, for Italian; van Berkum 1997, for Dutch). 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, 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 used an auditory naming task to examine how early English-French bilinguals and late English-French bilinguals react to grammatical gender violations when processing French DPs. They found that the early bilinguals showed clear gender priming effects, involving both facilitation and inhibition, whereas late bilinguals were 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 argued 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.

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 grammatical agreement. In this paper, we propose a model for grammatical agreement processing, building on Julien's (2005) DP model and incorporating the agreement mechanism developed in the Minimalist Program (Chomsky 1995). It is assumed that syntactic computation takes place during the syntactic formation of an utterance and during its parsing as well (Fong 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. First the correct components for assembly are laid out. The components include D with [*u*GEN], [*u*NUM], and [*u*DEF] features, and *hus* with the corresponding valued 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 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 (2) below:

(2) Probe-Goal account of agreement between D and N in building *det gamle huset* 'the old house'

PROBE		GOAL				
D	….	N		D	….	N
[*u*GEN]		[+neut]	Agree	~~[*u*GEN]~~		[+neut]
[*u*NUN]		[-pl]	⟶	~~[*u*NUN]~~		[-pl]
[*u*DEF]		[+def]		~~[*u*DEF]~~		[+def]
				↓		↓
				det		hus-*et*

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 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 a nice account of priming effects**.** 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. 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.

4. The experiment

4.1. Participants

Sixteen L2 participants and fourteen native speakers of Norwegian were recruited. The L2 speakers were selected from typologically different L1 backgrounds: Romance (5 informants), English (6 informants), and Chinese (5 informants). They had an average length of residence in Norway up to 14.8 years, and had achieved nativelike scores on an independent measure of proficiency in Norwegian (≥ 46 out of 50). So they were assumed to be endstate L2 speakers. The L1 participants matched in age and education backgrounds with their L2 counterparts. Both groups were naive about the goal of the study.

4.2. 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.

With respect to gender-marking, we assumed the two-way gender system of Norwegian, neuter gender and common 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. 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 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 were either singular (in which case the determiners matched with the noun targets in number), or in plural (in which case the determiners mismatched with the noun targets in number).

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 (note that gender information is not visible on D or N in plural). As Norwegian does not have an indefinite plural prenominal determiner comparable to the definite plural prenominal determiner (*de*), we kept the prenominal 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. Table 2 shows sample stimuli in gender-, number- and definiteness-marking (with the mismatched feature in bold).

Table 2. Sample stimuli in gender-, number-, and definiteness-marking

	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. 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. Care had been taken to counterbalance the word frequency (based on the frequency list of *the Oslo Corpus of Tagged Norwegian Text*) and the syllable length of nouns in discordant and concordant conditions for all the three feature groups. 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.

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.

4.3. 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.

4.4. 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 (ms). This result thus indicate that the RTs measured by the research assistant are reliable.

5. Results

5.1. 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).

5.2. The L1 participants

Table 3. Control group's Mean RTs to gender, number, and definiteness markings in each condition (RT in ms; standard deviation in brackets)

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

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

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.

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 [$t_{gen}(33)=2.91$, $p=.006$; $t_{num}(34)=5.21$, $p<.001$; $t_{def}(34)=2.86$, $p=.007$]. 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 based on the agreement process model thus borne out.

5.3. The L2 participants

The data analysis procedure for the L2 participants is basically the same as we used for L1 participants. 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 4 shows the mean RT in each cell and the average naming time difference in each agreement feature groups.

Table 4. The L2 group's mean RTs to gender, number, and definiteness markings in each condition (RT in ms; standard deviation in brackets)

	Grammatical agreement features			
Determiners	gender	number	definiteness	total
Concordant	602 (54)	602 (67)	638 (45)	616 (56)
Discordant	611 (56)	619 (38)	671 (51)	638 (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, 17ms difference in number-marking, and the 33ms difference in definiteness-marking [t_{gen}(33)=.479, p=.635; t_{num}(20)=.710, p=.486; t_{def}(33)=2.08, p=.054], were not significant. These results indicate that the endstate L2 speakers of Norwegian were insensitive to gender-, number-, and definiteness-marking.

In order to investigate the L1 transfer effect in 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 show that concordance-discordance difference in neither agreement feature group is significant [t_{gen}(164)=.465, p=.642; t_{num}(230)=.669, p=.504]. So contrary to our predictions, the L2 participants show no nativelike sensitivity to agreement violations in L2 grammar, suggesting that L1 transfer does not occur in L2 processing.

6. Discussion

6.1. A summary of findings

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;
- The L2 speakers were overall insensitive to DP internal agreement violations in L2 grammar; there was no significant concordance-discordance RT difference in either gender-, number-, or definiteness-marking group;
- 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.

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

6.2. How does the information on D influence the word recognition process?

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. 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.

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, we have 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 that feature mismatch, resulting in the inhibitory processing effect. As we see, it is the uninterpretable features that drive syntactic computation and identify the probe and goal. Since the Norwegian D carries [*u*GEN], [*u*NUM], and [*u*DEF] features, the syntactic computation applies automatically in L1 processing. It is thus hard for native speakers of Norwegian to suppress the relevant information on D.

6.3. L1/L2 differences in processing

One of the important findings from this experiment is that endstate L2 speakers of Norwegian were found to react to grammatical violations in Norwegian distinctly from native speakers. While native speakers uniformly 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 grammatical knowledge about Norwegian.

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) proposed by Clahsen and Felser (2006), who claim that the representations adult L2 learners compute during processing contain less syntactic detail than those of native speakers. 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 information of target nouns instead.

Note, however, although shallow processing has been argued as the predominant route for L2 processing, nativelike (full) processing route is claimed to be used in processing local mismatches. Clahsen and Felser (2006, p.111) writes:

> …during L2 processing, learners compute grammatical representations that lack complex hierarchical structure and abstract configurationally determined elements such as movement traces, and that nativelike grammatical processing is restricted to “local” domains such as word segmentation or morphosyntactic agreement between closely adjacent constituents.

Clahsen and Felser makes crucial use of the local domain/non-local domain distinction in their theory, and claim that shallow processing route is used in non-local domain sentence processing, whereas nativelike (full) parsing route is available to L2 learners in local domains. The claim was made presumably based on findings from experimental studies, though no theoretical support for such a distinction was provided. According to their 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 an alternative processing route. Thus, the present results extend SSH to local domains. As feature checking necessarily involves hierarchical relations such as C-command, we see no theoretical ground for a distinction between local and non-local domains in L2 processing.

6.4. Is there an L1 transfer effect in 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 et al. 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 conjuncture, a methodological question arises as to what count as ‘similarity’ between L1 and L2. In case of Sabourin’ 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 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 in 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, but rather in line with SSH, is mainly guided by lexical-semantic information. 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.

7. Conclusion

The present study used an auditory naming technique to investigate how gender, number, and 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, indicating 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 we have proposed, which incorporates the agreement mechanism developed in recent work of the Minimalist Program.

The L2 speakers' overall insensitivity to grammatical agreement violations was interpreted as an indication that automatic syntactic computation was not invoked. Following Shallow Structure Hypothesis, we have argued that L2 processing is mainly guided by lexical-semantic information. Therefore L2 speakers were able to ignore the agreement cues on D. This observation also point 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 opposing Clahsen and Felser (2006), who attribute non-nativelike L2 processing to inadequacy of L2 grammar, we argue instead that L2 processing may be governed by different processing mechanisms. In addition, the present study have shown that shallow processing route can be used in processing local mismatches, thus extending SSH to local domains.

References

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

Bates, Elizabeth and Liu, Hua (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.

Boersma, Paul & Weenink, David (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.

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

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.

Gillon-Dowens Margaret, Barber Horacio, Bergara Marta & Carreiras Manuel (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.

Grosjean, François, Dommergues, J.-Y., Cornu, E., Guillelmon, Delphine and Besson, C. (1994) The gender-marking effect in spoken word recognition. *Perception and Psychophysics* 56, 590-598.

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

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

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

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

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

Tokowicz, Natasha & MacWhinney, Brian (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.

van Berkum, Jos (1996) *The Psycholinguistics of Grammatical Gender*. Nijmegen: Nijmegen University Press.

Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)

edited by Jean Crawford,
Koichi Otaki, and Masahiko Takahashi

Cascadilla Proceedings Project Somerville, MA 2009

Copyright information

ISBN 978-1-57473-436-2 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. 2009. Sensitivity to DP-Internal Agreement Violations in L2 Grammar. In *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)*, ed. Jean Crawford, Koichi Otaki, and Masahiko Takahashi, 116-126. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2312.