Björn Lundquist*, Yulia Rodina, Irina A. Sekerina and Marit Westergaard

Gender Change in Norwegian Dialects: Comprehension is affected before Production

DOI 10.1515/lingvan-2016-0026
Received April 20, 2016; accepted August 7, 2016

Abstract: This article investigates language variation and change in the grammatical gender system of Norwegian, where feminine gender agreement is in the process of disappearing in some Northern Norwegian dialects. Speakers of the Tromsø ($N = 46$) and Sortland ($N = 54$) dialects participated in a Visual Word experiment. The task examined whether they used indefinite articles (*en*, *ei*, *et*) predictively to identify nouns during spoken-word recognition, and whether they produced feminine articles in an elicited production task. Results show that all speakers used the neuter indefinite article *et* as a predictive cue, but no speakers used the feminine *ei* predictively, regardless of whether they produced it or not. The masculine article *en* was used predictively only by the speakers who did not produce feminine gender forms. We hypothesize that in dialects where the feminine gender is disappearing, this change in the gender system affects comprehension first, even before speakers stop producing the feminine indefinite article.

Keywords: Psycholinguistics, variation, grammatical gender, language change, Northern Norwegian dialects, Visual World paradigm, PsychLingVar

1 Introduction

Currently, the Norwegian grammatical gender system is undergoing a change from three genders (masculine, feminine, neuter) to two (masculine, or common, and neuter). The change takes place across a number of dialects: Some dialects have already developed a stable two-gender system, while other dialects still show morphological distinctions between masculine and feminine. The Bergen dialect underwent the change centuries ago (e. g., Jahr 1998), while the Oslo dialect has evolved into a two-way gender system only recently (Lødrup 2011). In the Northern dialect of Tromsø, the distinction between feminine and masculine is in the process of disappearing. The change is characterized by substitution of the feminine indefinite article *ei* (e. g., *ei bok* 'a book') with the masculine form *en* (e. g., *en bok)*. There is considerable inter- and intra-speaker variation, especially in the younger generations (Rodina and Westergaard 2015). There is no spoken standard in Norway, and people generally speak their dialects in all situations. In contrast, there are two written standards, *bokmål* and *nynorsk*. Both have a three-gender system with distinctions between masculine, feminine, and neuter, but the *bokmål* standard, which is used predominantly across the majority of regions of Norway, also allows a two-gender system consisting of common gender (used for all masculine and feminine nouns) and neuter. Variation is thus present both in the spoken and written input, but the amount of variation differs from dialect to dialect.

Psycholinguistic research has provided evidence that monolingual speakers make use of gender markers and other inflectional categories to predict upcoming linguistic material (Grüter et al. 2012; Hopp 2013; Lew-Williams and Fernald 2007). Several studies have directly targeted the relation between production and comprehension of grammatical gender in the course of L1 and L2 language acquisition (Dussias et al. 2013; Grüter et al. 2012; Hopp 2013, 2016). The present study offers a novel take on

***Corresponding author: Björn Lundquist,** Department of Language and Culture, UiT The Arctic University of Norway, Postbox 6050 Langnes, 9037 Tromsø, Norway, E-mail: bjorn.lundquist@uit.no
Yulia Rodina, Department of Language and Culture, UiT The Arctic University of Norway, 9037 Tromsø, Norway
Irina A. Sekerina, Department of Psychology, College of Staten Island, 2800 Victory Blvd. 4S-108, Staten Island, New York 10314, USA
http://orcid.org/0000-0003-3859-3000
Marit Westergaard, Department of Language and Culture, UiT The Arctic University of Norway, 9037 Tromsø, Norway and NTNU Norwegian University of Science and Technology, Trondheim, Norway

production-comprehension interaction by investigating gender marking in production and comprehension in the context of language variation and change. Specifically, our aim is to investigate how gender marking on indefinite articles (*en, ei, et* for masculine, feminine and neuter gender, respectively) facilitates comprehension in Norwegian speakers who use either a two- or a three-gender system in their production. Do the speakers who do not use feminine *ei* in production nevertheless rely on it to predict upcoming nouns? Are speakers who encounter inconsistency in the gender system within the language community able to use the gender markers in real-time comprehension?

To address these questions, we conducted a Visual World Paradigm experiment (VWP) that targets noun anticipation effects triggered by the three indefinite articles in three groups of speakers in Northern Norway: One group has developed a consistent two-gender system, while the other two groups still have a consistent three-gender system. The two three-gender groups differ in the amount of variation they are exposed to in the relevant linguistic community: The first group (Sortland dialect) is situated in a fairly stable linguistic community whereas the other one (three-gender Tromsø dialect) is situated in a linguistic context with considerable contact with two-gender dialects. Given the differences between these dialects, we can test if the amount of variation within a speech community with respect to gender markers affects how these markers are used in real-time comprehension. The present study can thus provide important evidence on how comprehension is affected in relation to production during language change. More specifically, the study can give us insight into which of the following three hypotheses best capture the relation between comprehension and production in the context of language change:

1. Production and comprehension are equally affected: If gender distinctions are made in production, they are also used predictively in real-time comprehension.
2. Production is affected before comprehension: Speakers are still sensitive to three-gender distinctions made by other speakers in real-time comprehension, without necessarily making the same distinction in their own production.
3. Comprehension is affected before production: Speakers make three-gender distinctions in their production but do not use them to anticipate upcoming language, possibly due to inconsistency in the input.

The three hypotheses make different predictions about how the three groups of speakers will perform in a VWP experiment. Hypothesis 1 predicts that speakers who have a stable three-gender system will be able to use all three articles to predict an upcoming target noun, while two-gender speakers will only make use of *en* (masculine) and *et* (neuter). Hypothesis 2 predicts that all three groups will use all three articles to predict an upcoming target. Hypothesis 3 predicts that for masculine and feminine articles, the size of the effect should correlate with the stability of the gender system within the given group. The speakers of the three-gender system will make fewer distinctions in their comprehension than in their production: The inconsistency in the input will make the gender markers that are affected by the change less reliable predictors of upcoming linguistic material. It is less clear if speakers who have developed a stable two-gender system will also be affected by the inconsistency in the input, or if the inconsistency will simply be ignored by these speakers. For the neuter article, Hypothesis 3 predicts that both two- and three- gender speakers will use the gender information provided by neuter *et* to predict upcoming nouns, since the neuter gender is stable across all dialects.

2 The Norwegian gender system and the loss of feminine

We define grammatical gender according to a traditional view, expressed by Hockett (1958 253) as 'classes of nouns reflected in the behavior of associated words' (see also Corbett 1991).[1] In Table 1, we provide a brief

1 This means that we consider the definite suffixed forms (e. g., *hest-en* 'the horse', *seng-a* 'the bed', *hus-et* 'the house') not to be exponents of gender, but of declension class (Enger 2004; Lødrup 2011; Rodina and Westergaard 2015) and thus beyond the scope of this study.

Table 1: The traditional three-gender system of Norwegian.

Gender/SG	Masculine	Feminine	Neuter
Indefinite	*en hest* 'a horse'	*ei seng* 'a bed'	*et hus* 'a house'
Adjective	*en fin hest* 'a nice horse'	*ei fin seng* 'a nice bed'	*et fint hus* 'a nice house'
Double definite	*den hest-en* 'that horse.DEF'	*den seng-a* 'that bed.DEF'	*det hus-et* 'that house.DEF'

overview of the most important exponents of gender in the traditional three-way gender system, where masculine is considered to be a default (Trosterud 2001). Gender is marked in the singular on indefinite articles and to some extent on adjectives and prenominal determiners in double definite forms (e. g., *den hesten* 'that horse'). Indefinite articles (as well as possessive determiners, not discussed in this article) have distinct forms in all three genders, while virtually all adjectives and prenominal determiners in double definites distinguish only between common and neuter.

Gender assignment in Norwegian is relatively non-transparent. Nevertheless, there is a class of disyllabic (weak) nouns, characterized by the ending *-e* (e. g., *ei dukke* 'a doll'), which has been argued to be a reliable cue for feminine (Trosterud 2001). In dialects spoken in and around Tromsø, (i. e., Troms county), the ending *-a* has traditionally been used for weak feminines (e. g., *ei dukka),* distinguishing this class of nouns from masculines ending in *-e* (e. g., *en pinne* 'a stick'). Rodina and Westergaard (2015) report that there has been a change in the Tromsø dialect from the *-a* to the standard *-e* ending, which has preceded the change in the feminine gender forms, as even older speakers who consistently produce the feminine indefinite article *ei* use the traditional ending *-a* in weak feminines only 20 % of the time. This change from *-a* to *-e* may have blurred the distinction between masculine nouns and some feminines and may have been a contributing factor to the loss of feminine.

Data from a corpus of eight speakers of the Tromsø dialect (Anderssen 2006; Rodina and Westergaard 2015) show that the masculine indefinite article *en* is by far the most frequent one (62.6 %), whereas feminine *ei* and neuter *et* are equally infrequent, 18.9 % and 18.5 %, respectively. Moreover, Rodina and Westergaard (2015) showed that there was a striking asymmetry in the use of feminine *ei* by age: In contrast to adults (31–64 years), who use feminine *ei* 99 % of the time, adolescents (18–19 years) use it 56 %, and children (3–12 years) only between 7 % and 15 %. While the cause of the change is presumably sociolinguistic, Rodina and Westergaard (2015) argue that the nature of the change is driven by language acquisition, the contributing factors being syncretism, frequency, lack of transparency and a distinction between gender and declension forms.

3 The visual world paradigm and the processing of gender markers

Psycholinguistic research from the last 15–20 years has provided compelling evidence that speakers make use of the information provided by inflectional categories like case, number and gender agreement to predict upcoming linguistic material. Evidence comes from both ERP (Wicha et al. 2004; van Berkum et al. 2005) and VWP eye-tracking studies (Allopenna et al. 1998; Dahan et al. 2000; Lew-Williams and Fernald 2007). Several VWP studies have investigated how speakers use grammatical gender information present on articles and adjectives to rapidly locate a target object in a visual context. In the simple version of this paradigm, a participant is presented with two or more object pictures on a screen, and a spoken instruction like *Look at the (adjective) NOUN*, where the article (and possibly also the adjective) carries gender information. The participant's eye movements reflect whether the presence of gender markers facilitates looks to objects that match the grammatical gender of the article, i. e., if the participant uses the gender information to predict the target noun.

Although predictive effects triggered by gender information on articles and adjectives have been observed in many languages, e. g., French (Dahan et al. 2000), Spanish (Grüter et al. 2012; Lew-Williams and Fernald 2007), Russian (Sekerina and Trueswell 2011), German (Hopp 2016), and Dutch (Loerts et al. 2013), this effect can be modulated by the population tested, the morpho-syntactic system of the language, and the experimental set-up. Children and older adults, speakers with low working memory capacity, and L2 learners show smaller or no signs of anticipation of upcoming language compared to monolingual young controls (Huettig and Mani 2016). Furthermore, morpho-syntactic properties of the language also determine whether an anticipation effect is found. Loerts et al. (2013) found anticipation effects for the masculine gender in Dutch, but not for the neuter, and related this finding to the fact that the neuter gender is used to form all diminutives in Dutch, regardless of lexically specified gender. Since all nouns can in principle occur with a diminutive suffix, the neuter gender does not facilitate anticipation of upcoming nouns. Finally, the experimental set-up has been shown to highly affect the observed anticipation of upcoming language: Huettig and Guerra (2015) found that their participants used gender cues in Dutch predictively when sentences were presented at a slow speech rate. The effect was also found when the sentences were presented at a normal speech rate, but only if the participants had at least a four-second preview of the objects in the visual display. However, the length of preview did not matter when sentences were presented at a slow speech rate.

It is still an open question whether the lack of anticipatory looks in certain populations, morpho-syntactic systems, and experimental set-ups should be interpreted as a complete lack of predictive processing or whether the methodology is not fine-tuned enough to pick up the predictive signal (see Huettig and Mani 2016 for discussion). However, these factors are not relevant for our study, since all our participants are native adult speakers of Norwegian, and not only did they have a four-second preview of the objects, but the objects were also explicitly named for them. The morpho-syntactic system of Norwegian is balanced, with the neuter and the feminine equally frequent, so there is no *a priori* reason to expect that these two genders should have different effects on anticipation in a stable language situation.

3.1 Participants

This study was conducted at two locations in Northern Norway: The University of Tromsø ($N = 46$), and a high school in the town of Sortland ($N = 54$), 150 kilometers south of Tromsø. Tromsø was chosen because it is a location where the change has recently been attested, and where we know that speakers are exposed to considerable linguistic variation. Sortland, on the other hand, is in a more remote area and thus less exposed to dialect variation. The Tromsø participants ($N = 46$) consisted of students and staff at the University of Tromsø, varying in age (19–60) and dialectal backgrounds. The Sortland participants ($N = 54$) were high school students (age 16–18), all speaking the same local dialect.

Participants filled in a background questionnaire and a consent form. The Tromsø participants were rewarded with a gift card for the local coffee bar, and the Sortland participants received 40 NOK (5 USD) each, which was given to a shared school/class account. The study lasted on average 10 minutes.

3.2 The elicited production task

A short elicited production task was carried out together with a VWP eye-tracking experiment. The goal was to test the use of the feminine *ei* as well as the choice of final vowel (*-e*/*-a*) on weak feminines. Six production items appeared in groups of two after every six or seven comprehension items. We used the following procedure:

1. Two images appeared on the screen (e. g., a doll and a table), accompanied by a prerecorded voice: *Æ ser et bord. Ka ser du?* 'I see a table. What do you see?'

2. The participant replied, either by naming the object (*ei dukke/dukka* 'a doll') or with a full sentence (*Æ ser ei dukke/dukka*. 'I see a doll.').
3. One of the objects disappeared from the screen followed by the pre-recorded voice asking *Ka va det som forsvant*? 'What disappeared?'
4. The participant replied by naming the object in its definite form (*dukka* 'the doll') or a full sentence (*Dukka forsvant*. 'The doll disappeared').

The visual stimuli were images that had appeared as fillers earlier in the comprehension task. In total, participants had to name five indefinite and five definite feminine nouns. Three of the indefinite nouns were weak feminines, ending in *-e* or *-a* depending on dialect. The participants' responses allowed us to estimate their use of the feminine indefinite article *ei* and the choice of the final vowel *-e/-a*.[2]

The production task revealed that the majority of the Tromsø participants (33/46) consistently used feminine *ei* (Figure 1(A), left panels). Eleven participants used masculine *en* for the feminine nouns, and these speakers were not native speakers of the Tromsø dialect. Only two participants displayed inconsistent use of the articles (2–3 out of 5 items), and these two participants were excluded from the analysis. Most of the participants used *-e* as the final vowel in the weak feminine nouns (e. g., *ei/en dukke* 'a doll') (Figure 1(A), right panels). Based on the results from the production task, we split the Tromsø participants into two groups: Fem users ($N = 33$) and No-Fem users ($N = 11$).

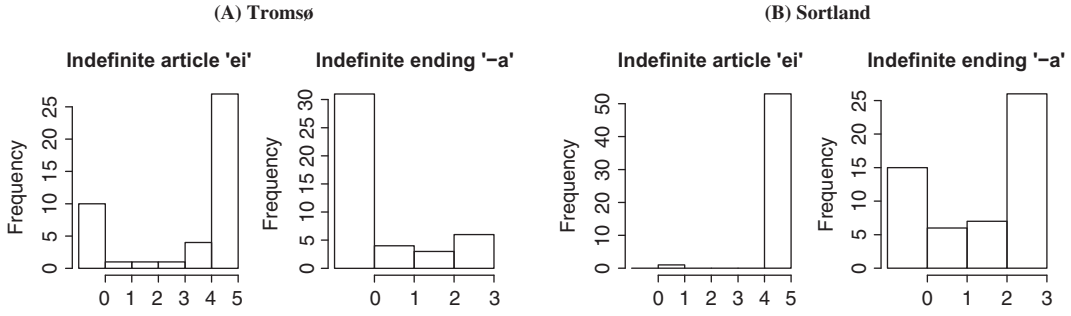

Figure 1: The use of feminine *ei* (left panels) and the ending *-e/-a* with weak feminines (right panels). (A) Tromsø (B) Sortland.

As expected, the participants in Sortland (Figure 1(B)) showed no variation in their production results: All but one participant (53/54) used the feminine *ei* consistently. This participant was excluded from further analysis. The rest of the group showed considerable variation with respect to the use of the *-e/-a* endings. In the analysis of the comprehension task, we thus focus on three groups of speakers: Fem users from Sortland ($N = 53$), Fem users from Tromsø ($N = 33$), and No-Fem speakers from Tromsø ($N = 11$).

3.3 The VWP eye-tracking comprehension experiment

The materials for our study were 64 experimental and four practice sets of picture stimuli. Each stimuli set consisted of four pictures: a target object (e. g., *bjorn* 'bear$_M$'), a competitor object of the same gender (e. g., *gris* 'pig$_M$') in the Same-Gender conditions, or a competitor object of a different gender (e. g., *esel* 'donkey$_N$') in the Different-Gender conditions, as well as two distractors (e. g., *kjøleskap*

2 As mentioned in footnote 1, the old feminine nouns have a different definite marker (*-a*) compared to masculines (*-en*), which has been treated as a difference in declension class. By making speakers produce definite forms as well as indefinite forms (step 4 in the procedure), we could check whether the declension system had been affected on par with the gender system. The results revealed that the declension split is still fully stable in Northern Norwegian.

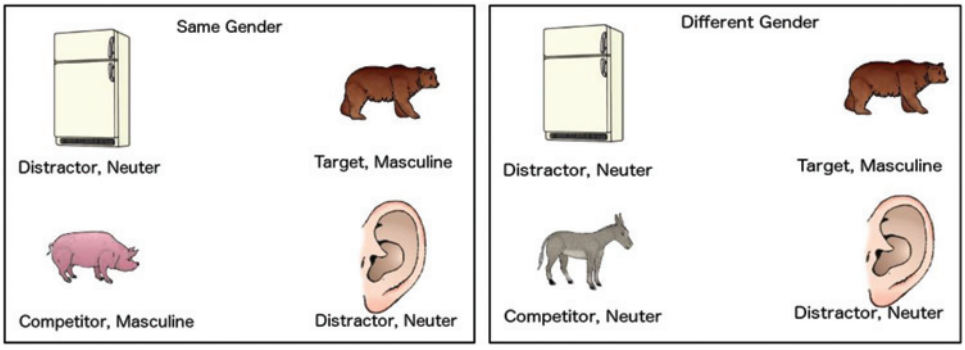

Figure 2: Examples of visual displays: Same-Gender (left panel) and Different-Gender condition (right panel).

'fridge$_N$' and *øre* 'ear$_N$') (see Figure 2). If the gender information is used to predict the target noun, participants should fixate the target object faster in the different gender condition compared to the same gender condition. The initial locations of the target and the gender competitor/distractor, as well as their destination locations, were counterbalanced across trials.

We manipulated the factors Gender Match (Same- or Different-Gender displays) and Gender of Target/Competitor (see Table 2). Due to restrictions imposed by length of the experiment and statistical power, we included only four of the six logically possible gender combinations: Neuter-Masculine (N-M), Masculine-Neuter (M-N), Masculine-Feminine (M-F) and Feminine-Masculine (F-M). The four conditions included were sufficient for testing the three hypotheses presented in the Introduction. If production and comprehension are affected simultaneously (Hypothesis 1), we expect to see an effect for the No-Fem speakers in the N-M and M-N conditions, and an effect in all conditions for the two groups of Fem speakers. If production is affected before comprehension (Hypothesis 2), we expect an effect for all three groups in all four gender conditions. If comprehension is affected before production (Hypothesis 3), the three conditions with masculine and feminine targets (M-N, M-F, F-M) could fail to show a same-different effect for the Fem speakers, especially for the Tromsø participants, where variation is abundant and change has been attested.

Table 2: Design of the VWP comprehension task.

Manipulation	Target	Competitor	Distractors
Same-Gender:			
C1: N-M	Neuter	Neuter	Masculine
C2: M-N	Masculine	Masculine	Neuter
C3: M-F	Masculine	Masculine	Feminine
C4: F-M	Feminine	Feminine	Masculine
Different-Gender:			
C1: N-M	Neuter	Masculine	Masculine
C2: M-N	Masculine	Neuter	Neuter
C3: M-F	Masculine	Feminine	Feminine
C4: F-M	Feminine	Masculine	Masculine

There were eight items per condition, with Gender Match manipulated as a between-participants and Gender of Target/Competitor as a within-participants factor. Participants were randomly assigned to one of two versions of the experiment, where the two versions had 16 Same- and 16 Different-Gender displays each, balanced between the four gender conditions.

The entire set consisted of 128 pictures, which were selected from Cycowicz et al. (1997). Half of the target and competitor pictures were masculine (i. e., two out of four gender conditions had masculine

targets and two out of four had masculine distractors), and the other half was equally split between feminine and neuter, which made both target and competitor/distractor items match the natural frequency of the different gender classes. The four gender conditions also functioned as filler items for each other.

Each trial started with a familiarization phase, in which the participants saw the visual display with the four pictures for 4 seconds while hearing a recording where the four items were named by a female speaker, recorded at 44,100 Hz. The objects were named in bare form, without a gender-marked article. After the familiarization phase, the cartoon figure Tweety appeared in the middle of the screen, saying *Æ gjemmer meg bak*... 'I am hiding behind... ' followed by an indefinite article and a noun. The four pictures reappeared at the onset of the article, and Tweety jumped into the target position 500 ms after the fixation of the target. For the Tromsø participants, the length of the article varied slightly between the different items, but for the Sortland participants, the duration of the article was always 550 ms.[3] For each group, the voice of the cartoon character was from a local dialect speaker of Northern Norwegian, recorded at 44,100 Hz.

The experiment was conducted with an SMI RED 500 eye-tracker, at a sampling rate of 250 Hz. The experiment was set up using the SMI's Experiment Center, and results were exported for analysis through the Proportions of looks module in BeGaze.

4 Results of the VWP comprehension experiment

In analyzing the results from the eye-tracking experiment, we fitted in total four linear mixed effects models using the lme4 package in R (Bates et al. 2015): One for each of the speaker groups, and one omnibus model including all three groups. Our dependent variable was the first fixation of the target, but the duration of the articles was removed because the articles differed in length. The predictors were gender (Neuter target-Masculine distractors, M-N, M-F and F-M) and condition (Same or Different). In the omnibus model, group was included as a fixed effect as well. The random effects included intercepts for Participant, Item and Order (placement on screen), and a by-participant random slope for condition (Same-Different). Since we were interested in the differences between the Same and Different conditions, we used contrasts which allowed precise comparisons of Same-Different distractors across conditions; see Table 4 in the Appendix for the contrast matrix. All trials where the target was not fixated before 1,500 ms after noun onset were removed from the analysis: In total 51 trials (1.4 % of the total number of trials), evenly spread across gender, conditions and groups. The full summaries of the models are presented in Tables 5–8 in the Appendix.

Average duration of the latencies to fixate the target and difference between Same- and Different-Gender conditions for the four gender conditions (N-M, M-N, M-F, F-M) and the three groups of participants (No-Fem Tromsø, Fem Tromsø, and Fem Sortland) are shown in Table 3. The lmerTest package (Kuznetsova et al. 2015) was used to obtain P-values from separate models for the three speaker groups.

All three groups located the target significantly faster in the Different condition compared to Same condition for the N-M gender: No-Fem: $\beta = -109$, SE(β) = 51.1, $t > 2$, Fem: $\beta = -118$, SE(β) 32.9, $t > 2$), Sortland: $\beta = -62$, SE(β) = 28.9, $t > 2$). However, only the No-Fem group showed a significant difference between the Same and Different conditions in the M-N condition: No-Fem: $\beta = -112$, SE(β) = 51.1, $t > 2$, Fem: $\beta = -16$, SE (β) = 32.9, ns, Sortland: $\beta = 21$, SE(β) = 28.9, ns.). None of the groups showed any effect in gender conditions 3 and 4. In the omnibus model, the Sortland group was the intercept. The result from this model confirmed the Same-Different effect for N-M in the Sortland group, and also showed that the Same-Different manipulation did not yield different results for N-M in the Fem and No-Fem groups compared to the Sortland group (see Table 4 in the Appendix). It further confirmed that the Same-Different manipulation was significantly different in M-N for the No-Fem group ($\beta = -133$, SE(β) = 66, $t > 2$) compared to the Sortland

3 The fact that no difference was observed between the three-gender speakers in Tromsø and the three-gender speakers in Sortland suggests that slight variation in the article length is unlikely to be a contributing factor.

Table 3: Mean latencies in milliseconds of looks to the target (Standard Errors in brackets).

	N-M	M-N	M-F	F-M
NO-FEM SPEAKERS (TROMSØ, N = 11):				
Different-Gender	474 (46)	402 (49)	550 (44)	444 (39)
Same-Gender	570 (48)	518 (35)	518 (40)	464 (42)
Difference	−96*	−116*	32	−20
FEM SPEAKERS (TROMSØ, N = 33):				
Different-Gender	420 (30)	515 (20)	547 (23)	479 (25)
Same-Gender	544 (25)	531 (20)	538 (23)	520 (22)
Difference	−124***	−16	11	−41
FEM SPEAKERS (SORTLAND, N = 53):				
Different-Gender	479 (19)	518 (20)	522 (19)	478 (20)
Same-Gender	549 (19)	515 (19)	560 (17)	497 (19)
Difference	−70*	3	−38	−19

Note: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

group. There were no other relevant effects. Moreover, the model also showed that there was no significant difference in the overall latencies for the three groups, indicating that the overall performance for the three groups were similar.

Figures 3–5 show the time courses of looks to the target. Each of the four plots presents looks to the target in the Same- vs. Different-Gender conditions. The looks to target in N-M when the target is neuter and the distractors are masculine diverge in the Different-Gender condition significantly earlier that in the Same-Gender condition. In the No-Fem group, this difference is masked by the high standard error because

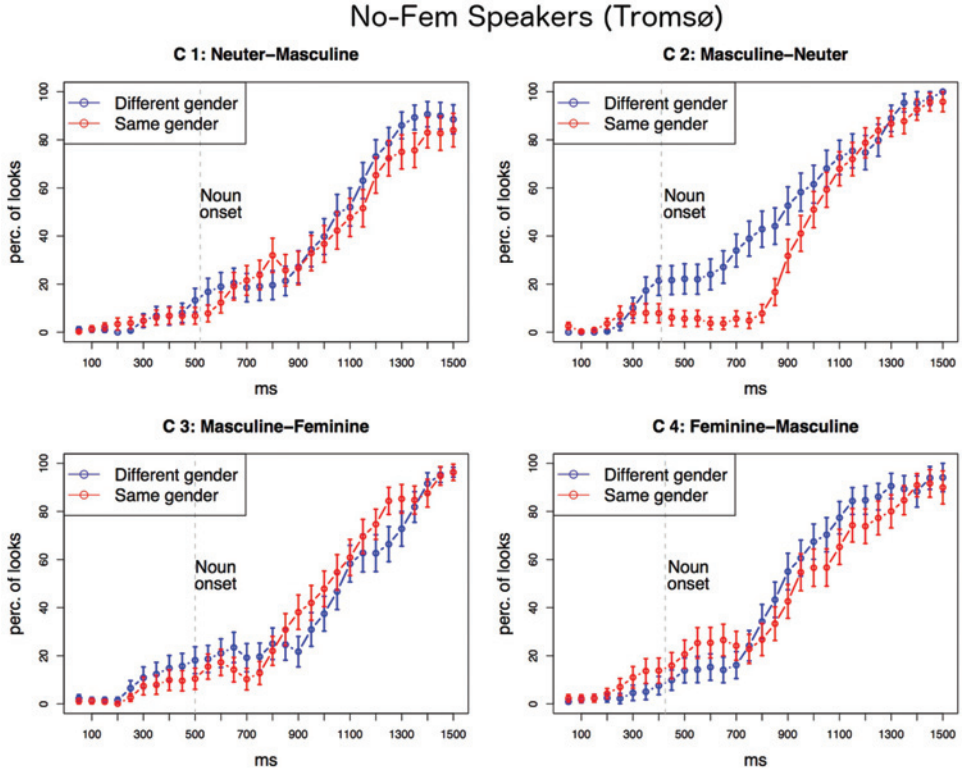

Figure 3 : Time course of No-Fem speakers' looks to target in Same- and Different-Gender trials. Error bars indicate 1 standard error above and below means.

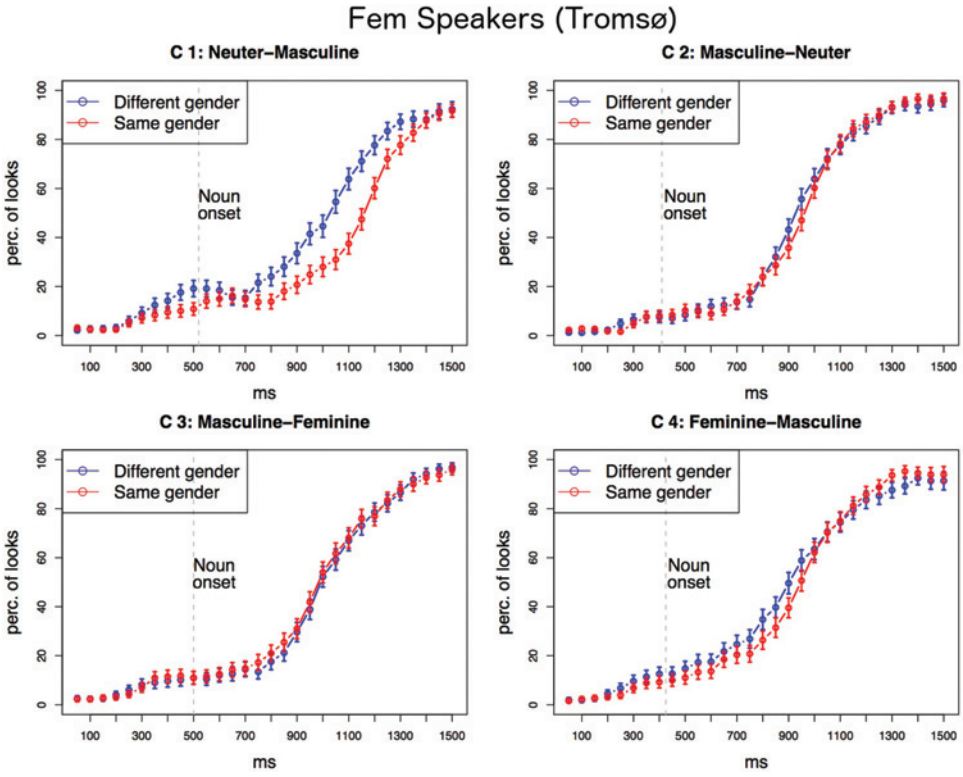

Figure 4: Time course of Tromsø Fem speakers' looks to target in Same- and Different-Gender trials. Error bars indicate 1 standard error below means.

of the small sample size (11). For the No-Fem speakers, the divergence of looks to the target between the Same- and Different-Gender conditions is also found in the M-N condition, but there was no difference in looks for the remaining gender conditions (M-F, F-M) in either of the groups. The time course plots are thus fully coherent with the results from the time to fixation results given by models 1–4.

5 Discussion

The present study investigated real-time comprehension of gender marking on indefinite articles (*en, ei, et* for masculine, feminine and neuter gender, respectively) in two dialect areas in Northern Norway, where a change in progress in the form of the loss of feminine gender has been attested. An experiment consisting of elicited production and a VWP eye-tracking task was carried out in two areas: Tromsø, a university city with considerable dialect mix, and Sortland, a remote and more stable dialect area. The adults in the Tromsø area were split into two groups based on their performance in the production task.

We observed an effect of anticipation for all three groups in the Neuter-Masculine condition, indicating that speakers of Norwegian make use of gender markers on indefinite articles to predict upcoming nouns. Otherwise, the two-gender speakers showed slightly different patterns compared to the three-gender speakers. The groups of Fem users (Tromsø and Sortland) showed no sign of anticipatory looks to either masculine or feminine targets. However, the No-Fem group used the masculine as a predictive cue, but only when the competitor belonged to the reliably distinct neuter gender class.

Thus, we observed a production-comprehension asymmetry: The two groups of Fem speakers made more contrasts in their production compared to their comprehension. Hypothesis 1 (comprehension and production equally affected) or Hypothesis 2 (production affected before comprehension) can thus not be correct.

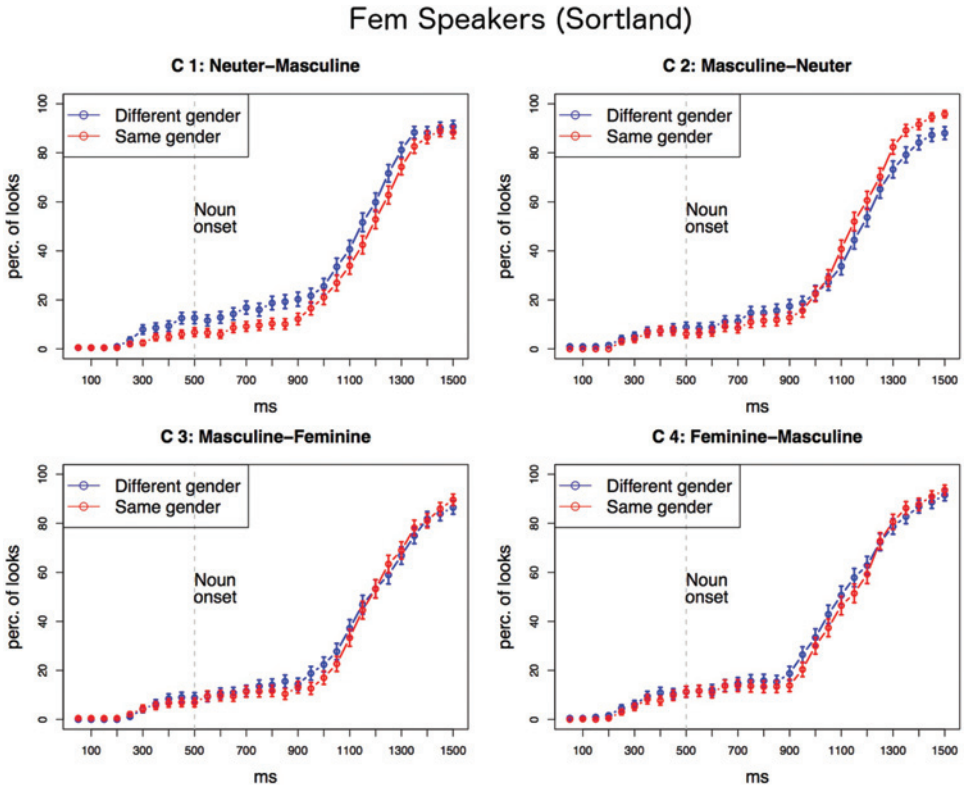

Figure 5: Time course of Sortland Fem speakers' looks to target in Same- and Different-Gender trials. Error bars indicate 1 standard error above and below means.

How can the absence of a predictive effect for the feminine *ei* be explained, in speakers who still produce it? One plausible hypothesis is that in language processing, comprehension is affected before production in a context of language change. In situations with considerable linguistic variation present, such as in Tromsø, speakers 'adjust' their comprehension by paying less attention to linguistic features that are not consistently used in the environment, before they change their own production. This could also explain the absence of anticipatory looks to target after the masculine article *en*, observed for the Fem speakers. Given that *en* can appear with either masculine or feminine nouns, the speakers may fail to perceive it as a stable predictor of upcoming linguistic material.

In contrast, the No-Fem speakers seem to be able to make use of the gender information present on *en* to predict the target noun. It may seem surprising that the *en-ei* inconsistency does not have a similar effect for the No-Fem speakers as it does for the Fem speakers, given that both groups encounter the *en-ei* variation in their speech community. As was pointed out in Section 3.2, the strict No-Fem speakers in our experiment were not native speakers of the Tromsø dialect. Rather, they were people who have moved to Tromsø from more stable two-gender areas, such as Oslo and Bergen. There may be several reasons why the presence of the feminine article does not interfere with the stable two-gender system. One possibility is that the phonological distinction between *ei* and *en* is so subtle to the non-native ear that it is simply not perceived. Alternatively, *ei* may be interpreted as a dialectal allomorph of *en*, with a semi-regular distribution. This could be tested in a follow-up experiment using either masculine or feminine target nouns preceded by the feminine article in a set-up with neuter distractors. If the speculations above are on the right track, feminine *ei* should be an equally strong predictor as *en* when the distractors are neuter, even when the target is actually a masculine noun.

We are aware that the group of No-Fem speakers was very small, and that the results from this group should be interpreted with caution. A follow-up experiment in a stable two-gender area will be conducted

in the near future, to confirm the results from the Tromsø No-Fem speakers. For now, we choose to interpret the current result from the No-Fem group as evidence of a complete loss of the feminine gender, both in production and comprehension, i. e. a fully developed two-gender system. In this system, the masculine *en* is a strong predictor, just like neuter *et*, whereas the feminine article is not. Thus, for these speakers, the absence of anticipatory looks to target in the condition where the target is masculine and distractors feminine is not surprising: The non-neuter nouns assume common gender, and the *en* article gives no further information than 'non-neuter'.

The results from the Sortland Fem speakers are perhaps the most surprising and interesting. We expected that the amount of variation of a given feature within a linguistic community would correlate with the extent to which the speakers in that community would use that feature in real-time processing. However, in Sortland we found no signs of variation in production in the use of gender-marked articles in the linguistic community, in contrast to Tromsø. One could perhaps take an extreme position and argue that feminine and masculine indefinite articles are never used predictively in stable three-gender dialects of Norwegian. However, we have preliminary results from a follow-up study from another Northern dialect (Finnsnes) with more reliable gender cues, showing that this is not the case. Furthermore, fully symmetric predictive power of gender agreement has been found in other three-gender languages (Hopp 2013 for German). We hypothesize that inconsistent use of gender markers in the input from outside of the immediate linguistic community, i. e., in mass media (TV, Internet, and other written sources) can facilitate, if not drive the restructuring of the three-gender system to a two-gender one. This suggests that speakers' grammatical system is malleable and can be radically affected by sociolinguistic factors as well as the primary linguistic input.

The result for the Fem speakers is in many ways counter-intuitive. It implies that speakers must have acquired a stable connection between feminine nouns and the article *ei* and masculine nouns and the article *en*. At the stage of acquisition, the distinction between *en* and *ei* must have been fully transparent in the speech stream for the learners, otherwise the distinction could not have been learned. At a later stage, after encountering variation from a wider range of input sources, the speakers seem to have stopped paying attention to *en* and *ei* in the input. Still, the early acquired gender distinctions continue to govern language production. The speakers have thus been attrited, but the attrition has so far only affected the comprehension system, and not production. It is not necessarily the case that the attrition in the comprehension system will affect production in the long term. In principle, it is possible that all dialect speakers become attrited when they encounter a sufficient amount of input from other dialects, but the attrition will not necessarily be noticed in production. As long as the production is not affected, we will not see language change.

Finally, it is possible that the loss of the feminine gender marker in Norwegian is affected by the use of *-e*/*-a* endings of the weak feminines. It has been known since the seminal work of Karmiloff-Smith (1979) that the shape of the noun is a very strong predictor for gender, especially in children. In the Tromsø dialect, where gender change has been documented, there is a preference for the *-e* ending for the weak feminines (Rodina and Westergaard 2015), while in Sortland, the traditional *-a* ending is used more frequently (Figure 2). The effect of the use of *-e*/*-a* could be a contributing factor for the loss of feminine gender.

To conclude, the evidence obtained in our study suggests that the change from a three- to a two-gender system is complete for the No-Fem speakers. For the Fem speakers, the change is presumably in progress, and the system is unstable, as suggested by the lack of a predictive effect for both masculine and feminine articles in spoken-word recognition. Thus, in language change, comprehension in the form of spoken-word recognition is affected early, whereas change in production is gradual and slower. This resembles the results from highly proficient L2 speakers who have a near-native production of gender markers, but fail to use them in predictive processing (see Grüter et al. 2012). For native speakers, similar asymmetries between production and comprehension have been observed in the context of phonological change. In near-merger situations, where a phonemic distinction is about to disappear, speakers continue to produce the distinction even though they are unable to hear it (Herold 1990; Labov 2001).[4]

4 We thank Meridith Tamminga for pointing out this phenomenon to us.

Finally, our current study provides predictions for further research. We hypothesize that speakers of dialects that still have a consistent split in the final vowel (*-e* for masculines, and *-a* for feminines) will have a less vulnerable gender system, and these speakers should be more likely to use all three articles in real-time comprehension. To investigate this, further evidence from speakers of other Norwegian dialects is needed, and this work is currently under way in our laboratory. There are further dialects in western Norway that mark the masculine-feminine distinction on adjectives as well articles. These dialects may also be less vulnerable to change, and the relation between production and comprehension of the three genders may be fully symmetric here. As was mentioned in the introduction, the two written standards differ in the specification of gender. The areas investigated in our study all use *Bokmål* as the written standard. A follow-up study could directly target an area where the more conservative *Nynorsk* (with obligatory masculine-feminine contrast) is used. Given the linguistic situation in Norway with two written standards and considerable dialectal variation, a series of follow-up studies could give us valuable insight into how factors like written language, morphological regularities and the sociolinguistic situation affect production and comprehension during language change.

Appendix

Table 4: Orthogonal contrasts used in the models below. The factors gender and condition have been collapsed into one factor with eight levels.

	N-M		M-N		M-F		F-M		F-N vs M-F	F vs N	M-N vs M-F
	Different	Same	Different	Same	Different	Same	Different	Same			
N-M Different	0.5		0		0		0		0.5	−0.5	0
N-M Same	−0.5		0		0		0		0.5	−0.5	0
M-N Different	0		0.5		0		0		−0.5	0	−0.5
M-N Same	0		−0.5		0		0		−0.5	0	−0.5
M-F Different	0		0		0.5		0		−0.5	0	0.5
M-F Same	0		0		−0.5		0		−0.5	0	0.5
F-M Different	0		0		0		0.5		0.5	0.5	0
F-M Same	0		0		0		−0.5		0.5	0.5	0

Table 5: Model coefficients (in ms) for average difference in the latencies to target fixation from noun onset in Same and Different conditions, all groups (3051 observations, 98 participants, 32 items, 4 orders (different placements on screen), 3 groups, 8 gender/conditions (see Table 4)). Random intercepts for Participant, Item and Order, and by-participant random slope for condition (Same-Different). The intercept is the overall average for the Sortland group, and the two following lines are the β coefficients for the overall means of the two Tromsø groups.

Fixed effect	Estimate	SD	*t*	Random effect		Variance
Intercept (Sortland)	544	19.8	27.5	Participant	Intercept	11125
Fem (Tromsø)	−33	22	−1.48	Participant	Same	3554
No-Fem (Tromsø)	−51	33.1	−1.54	Item	Intercept	4409
N-M Different-Same	−62	27.2	−2.27	Order	Intercept	275
M-N Different-Same	22	27.1	0.8			
M-F Different-Same	−22	27.3	−0.8			
F-M Different-Same	−25	27.2	−0.91			
F-N vs M	−37	26.9	−1.39			
F-M vs N-M	−52	38	−1.37			
M-F vs M-N	42	38	1.1			
Fem: N-M Different-Same	−64	44.1	−1.46			
Fem: M-N Different-Same	−35	44	−0.8			

(*continued*)

Table 5: (*continued*)

Fixed effect	Estimate	SD	*t*	Random effect	Variance
Fem: M-F Different-Same	32	44.4	0.72		
Fem: F-M Different-Same	−7	44.4	−0.16		
Fem: F-N vs M	−5	21	−0.24		
Fem: F-M vs N-M	67	30	2.26		
Fem: M-F vs M-N	−23	29.8	−0.78		
No-Fem: N-M Different-Same	−43	66.7	−0.65		
No-Fem: M-N Diferent-Same	−133	66.6	−2.0		
No-Fem: M-F Different-Same	53	65.5	0.82		
No-Fem: F-M Different-Same	4	67.3	0.05		
No-Fem: F-N vs M	23	31.7	0.71		
No-Fem: F-M vs N-M	−13	45.3	−0.29		
No-Fem: M-F vs M-N	26	44.6	0.58		

Table 6: Sortland, Model coefficients (in ms) for average difference in the latencies to target fixation from noun onset in same and different conditions (1,683 observations, 54 participants, 32 items, 4 orders (different placements on screen), 3 groups, 8 Gender/Conditions (see Table 4)). Random intercepts for Participant, Item and Order, and a by-participant random slope for condition (Same-Different). The intercept is the overall mean for the Sortland group. Degrees of freedom and *p*-values were obtained from the lmerTest package.

Fixed effect	Estimate	SD	df	*t*	*p*	Random effect		Variance
Intercept	543	18.13	62.20	30	<2e-16***	Participant	Intercept	11583
N-M Different-Same	−62	28.88	428.30	−2.157	0.0316*	Participant	Condition Same	4252
M-N Different-Same	21	28.76	424.00	0.731	0.4654	Item	Intercept	4122
M-F Different-Same	−21	28.96	429.10	−0.718	0.4731	Order	Intercept	0
F-M Different-Same	−25	28.85	427.10	−0.874	0.3828			
F-N vs M	−38	26.53	31.50	−1.444	0.159			
F-M vs N-M	−54	37.52	31.50	−1.448	0.157			
M-F vs M-N	41.55	37.52	31.50	1.107	0.276			

Table 7: Fem, Tromsø, Model coefficients (in ms) for average difference in the latencies to target fixation from noun onset in same and different conditions, (1,030 observations, 11 participants, 33 items, 4 orders (different placements on screen), 3 groups, 8 Gender/Conditions (see Table 4)). Random intercepts for Participant, Item and Order, and a by-participant random slope for condition (Same-Different). The intercept is the overall mean for the Tromsø Fem group. Degrees of freedom and *p*-values were obtained from the lmerTest package.

Fixed effect	Estimate	SD	df	*t*	*p*	Random effect		Variance
Intercept	512	25	6.89	21	1.9e-07***	Participant	Intercept	9971
N-M Different-Same	−118	32.93	214	−3.595	0.000402***	Participant	Condition Same	5713
M-N Different-Same	−16	32.86	212	−0.495	0.621404	Item	Intercept	7201
M-F Different-Same	1	33.31	222	0.023	0.981849	Order	Intercept	664
F-M Different-Same	−31	33.23	220	−0.924	0.356622			
F-N vs M	−41	33.6	28	−1.23	0.2297			
F-M vs N-M	18	47.6	28	0.37	0.7101			
M-F vs M-N	19	47.5	28	0.4	0.692			

Table 8: No-Fem, Tromsø, Model coefficients (in ms) for average difference in the latencies to target fixation from noun onset in same and different conditions, all groups (388 observations, 11 participants, 32 items, 4 orders (different placements on screen), 3 groups, 8 Gender/Conditions (see Table 4)). Random intercepts for Participant, Item and Order, and a by-participant random slope for condition (Same-Different). The intercept is the overall mean for the No-Fem group. Degrees of freedom and *p*-values were obtained from the lmerTest package.

Fixed effect	Estimate	SD	df	*t*	*p*	Random effect		Variance
Intercept	495	50.69	9.6	10	2.62e-06***	Participant	Intercept	17105
N-M Different-Same	−109	51.1	21	−2.128	0.0345*	Participant	Condition Same	892
M-N Different-Same	−112	51.1	211	−2.194	0.0293*	Item	Intercept	7770
M-F Different-Same	31	49.9	206	0.625	0.5330	Order	Intercept	3789
F-M Different-Same	−32	51.78	216	−0.615	0.5394			
F-N vs M	−14	40.1	27.80	−0.347	0.7314			
F-M vs N-M	−58	57.07	28.29	−1.017	0.318			
M-F vs M-N	65	56.33	27.31	1.154	0.258			

Acknowledgments: We thank Sølvi Grønning Riise and Anne-Margrethe Albertsen for fieldwork assistance and help with producing the experimental material, Martin Corley for initial help with the data analysis as well as the audience of CUNY 2016 for insightful comments. Special thanks to all our 98 participants and the staff at VGS Sortland (Sortland high school) for facilitating the fieldwork.

Funding: This work is supported by the Department of Language and Culture, UiT The Arctic University of Norway, the Research Council of Norway through its Centers of Excellence funding scheme, project number 223265 (the second author) and a 2016 research/teaching grant to Norway from the U.S. Fulbright Program (the third author).

References

Allopenna, Paul D., James S. Magnuson & Michael K. Tanenhaus. 1998. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory & Language* 38. 419–439.

Anderssen, Merete. 2006. *The acquisition of compositional definiteness in Norwegian*. University of Tromsø dissertation.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. ArXiv e-print; *Journal of Statistical Software*, http://arxiv.org/abs/1406.5823.

Corbett, Greville G. 1991. *Gender*. Cambridge University Press.

Cycowicz, Yael M., David Friedman, Mairav Rothstein & Joan Gay Snodgrass. 1997. Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology* 65. 171–237.

Dahan, Delphine, Daniel Swingley, Michael K. Tanenhaus & James S. Magnuson. 2000. Linguistic gender and spoken-word recognition in French. *Journal of Memory & Language* 42. 465–480.

Dussias, Paola E., Jorge R. Valdés Kroff, Rosa E. Guzzardo Tamargo & Chip Gerfen. 2013. When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition* 35. 353–387.

Enger, Hans-Olav. 2004. On the relation between gender and declension: A diachronic perspective from Norwegian. *Studies in Language* 28(1). 51–82.

Grüter, Theres, Casey Lew-Williams & Anne Fernald. 2012. Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research* 28. 191–215.

Herold, Ruth. 1990. *Mechanisms of merger: The implementation and distribution of the low back merger in eastern Pennsylvania*. University of Pennsylvania dissertation.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: MacMillan.

Hopp, Holger. 2013. Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research* 29(1). 33–56.

Hopp, Holger. 2016. Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research* 32(2). 277–307.

Huettig, Falk & Ernesto Guerra. 2015. Testing the limits of prediction in language processing: Prediction occurs but far from always. *21st Annual Conference on Architectures and Mechanisms for Language Processing* (AMLaP 2015), Malta.

Huettig, Falk & Nivedita Mani. 2016. Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience* 31(1). 19–31.

Jahr, Ernst-Håkon. 1998. Kontakten nedertysk– nordisk i hansatida: mer enn bare lån? [Language contact between Low German and Nordic languages during the Hansa period: More than just borrowing?]. *Norsklæreren* 1. 34–43.

Karmiloff-Smith, Annette. 1979. *A functional approach to child language: A study of determiners and reference*. Cambridge: Cambridge University Press.

Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen. 2015. lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-29. http://CRAN.R-project.org/package=lmerTest.

Labov, William. 2001. *Principles of linguistic change, volume II, social factors*. Oxford: Wiley-Blackwell.

Lew-Williams, Casey & Anne Fernald. 2007. Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science* 18(3). 193–198.

Loerts, Hanneke, Martijn Wieling & Monika S. Schmid. 2013. Neuter is not common in Dutch: Eye movements reveal asymmetrical gender processing. *Journal of Psycholinguistic Research* 42(6). 551–570.

Lødrup, Helge. 2011. Hvor mange genus er det i Oslo-dialekten? [How many genders are there in the Oslo dialect?]. *Maal og Minne* 2. 120–36.

Rodina, Yulia & Marit Westergaard. 2015. Grammatical gender in Norwegian: Language acquisition and language change. *Journal of Germanic Linguistics* 27. 145–187.

Sekerina, Irina A. & John C. Trueswell. 2011. Interactive processing of contrastive expressions by Russian children. *First Language* 31(1–2). 61–87.

Trosterud, Trond. 2001. Genustilordning i norsk er regelstyrt. [Gender assignment in Norwegian is rule-based]. *Norsk Lingvistisk Tidsskrift* 19. 29–57.

Van Berkum, Jos J. A., Colin M. Brown, Pienie Zwitserlood, Valesca Kooijman & Peter Hagoort. 2005. Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology*. Learning, Memory, and Cognition 31. 443–467.

Wicha, Nicole Y. Y., Eva M. Moreno & Marta Kutas. 2004. Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience* 16. 1272–1288.