LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

EMPIRICAL STUDY

Bilectal Exposure Modulates Neural Signatures to Conflicting Grammatical Properties: Norway as a Natural Laboratory

Maki Kubota,[a] Jorge González Alonso,[a,b] Merete Anderssen,[a] Isabel Nadine Jensen,[a] Alicia Luque,[a,b] Sergio Miguel Pereira Soares,[c] Yanina Prystauka,[a] Øystein A. Vangsnes,[a,d] Jade Jørgen Sandstedt,[a,e] and Jason Rothman[a,b]

[a]UiT The Arctic University of Norway [b]Nebrija University [c]Max Planck Institute for Psycholinguistics [d]Western Norway University of Applied Sciences [e]Volda University College

Abstract: The current study investigated gender (control) and number (target) agreement processing in Northern and non-Northern Norwegians living in Northern Norway.

CRediT author statement – The first two authors contributed equally to this work. **Maki Kubota**: investigation; formal analysis; visualization; data curation; writing – original draft preparation; writing – review and editing. **Jorge González Alonso**: conceptualization; methodology; investigation; writing – original draft preparation; writing – review and editing. **Merete Anderssen**: conceptualization; methodology; funding acquisition; writing – review and editing. **Isabel Nadine Jensen**: methodology; investigation. **Alicia Luque**: investigation. **Sergio Miguel Pereira Soares**: investigation. **Yanina Prystauka**: investigation. **Øystein A. Vangsnes**: funding acquisition, writing – review and editing. **Jade Jørgen Sandstedt**: formal analysis; writing – review and editing. **Jason Rothman**: conceptualization; methodology; funding acquisition; original draft preparation; writing – review and editing; project administration.

A one-page Accessible Summary of this article in nontechnical language is freely available in the Supporting Information online and at https://oasis-database.org

This project was funded by AcqVA Aurora Center, UiT The Arctic University of Norway. Maki Kubota, Yanina Prystauka, Alicia Luque and Jason Rothman are partly supported by the HeLPiNG grant from the TromsØ Forskningsstiftelse (TromsØ Research Foundation). Jorge González Alonso is partly supported by grant 2020-T1/HUM-20037 from the Community of Madrid and project 326487-YRT from the Research Council of Norway.

Correspondence concerning this article should be addressed to Maki Kubota, Department of Language and Culture, The Arctic University of Norway, Tromsø, Norway. Email: maki.kubota@uit.no

The handling editor for this manuscript was Sarah Grey.

DOI: 10.1111/lang.12608

Participants varied in exposure to Northern Norwegian (NN) dialect(s), where number marking differs from most other Norwegian dialects. In a comprehension task involving reading NN dialect writing, P600 effects for number agreement were significantly affected by NN exposure. The more exposure the NN nonnatives had, the larger the P600 was, driven by the *presence* of number agreement (ungrammatical in NN). In contrast, less exposure correlated to the inverse: P600 driven by the *absence* of number agreement (ungrammatical in most other dialects). The NN natives showed P600 driven by the *presence* of number agreement regardless of exposure. These findings suggests that bilectalism entails the representation of distinct mental grammars for each dialect. However, like all instances of bilingualism, bilectalism exists on a continuum whereby linguistic processing is modulated by linguistic experience.

Keywords bilectalism; ERP; syntactic processing; linguistic experience

Introduction

Bilectalism (also sometimes referred to in the literature as "bidialectalism") refers to a case where individuals—bilectals—are sufficiently exposed to two distinct varieties (e.g., two spoken dialects) of what is largely described under a singular linguistic label, for example, Norwegian. Crucially, the grammars of these varieties must be sufficiently different, even if the varieties are mutually intelligible, so one would expect some level of competing underlying grammatical representations (Chambers & Trudgill, 1998).

Although there may be valid reasons to question the treatment of bilectalism as a subcategory of bilingualism, the relevant question is whether empirical facts warrant such a label. Investigating the way in which closely related varieties are acquired and processed, including how they may affect one another, should play a significant role in answering this question. As yet, little is known about how potential bilectals process grammatical information in distinct dialects to which they become (increasingly) exposed (e.g., Southern Norwegians who have moved to a new dialect region in Northern Norway later in life). Even less is known about how the brain processes linguistic information in bilectal contexts. In the present paper we endeavor to address these important gaps in the literature.

Here we investigate bilectal processing among individuals living in Northern Norway, who vary considerably in terms of their experience with Northern Norwegian (NN) dialects (including both native[1] Northerners and speakers of other dialects who relocated to the North). Specifically, we test how they process grammatical properties in NN, both properties (e.g., gender agreement) that are shared across Norwegian regional dialects (as well as both standardized written varieties, Bokmål and Nynorsk) and a property that is unique to NN (namely, number (non)agreement). Both properties (gender and number (non)agreement) are categorical (nonoptional) in each dialect. Unlike previous studies that examined dialectal processing of one specific syntactic property (e.g., Norwegian gender: Lundquist & Vangsnes, 2018), our study manipulated the degree of configurational dialectal alignment by pairing one property where dialect grammars coincide (gender agreement) with one where NN is unique compared to other regional varieties and the written standards (number agreement). Specifically, NN prohibits predicative number agreement, whereas nearly all other spoken and written varieties of Norwegian require it (Sandøy, 1988). Stimuli were presented in NN *dialektskriving* "dialect writing" or *talemålsnær skriving* "speech-close writing": a nonstandardized written form of NN dialect with a high incidence of local or regional dialect features, frequently used by NN dialect users in digital social media (Røyneland & Vangsnes, 2020; see below for further description of the Norwegian language context and the grammatical conditions tested). By correlating bilectals' event-related potential responses to their degree of experience with the target dialect (NN), while controlling for native speaker status, we aim to elucidate linguistic processing patterns arising in relation to the gradient spectrum of experiences with multiple grammars, specifically within the context of bilectalism.

Background Literature

Electrophysiological Measures of Agreement Processing in Second Language Literature

The present study uses event-related potentials (ERPs), a technique in which visual or auditory stimuli are presented concurrently with the recording of a subject's electroencephalogram (EEG), time-locking the signal to the presentation of a certain stimulus (or event) of interest. Investigating fluctuations in the EEG signal that unfold during the processing of a linguistic stimulus now has a decades-long tradition in psycholinguistics and has provided insights into some of the most fundamental aspects of language processing. One such feature of real-time language use is the computation of syntactic relationships between words at a distance. Long-distance dependencies are essential to language, but they represent a challenge during incremental sentence processing because a linguistic element that demands some kind of correspondence (e.g., a syntactic subject) must be held in memory until a subsequent, adequate element (e.g., a verb) provides the missing information. One prominent example of these long-distance dependencies is morphosyntactic agreement, where words that are meant to be interpreted together in some way receive systematically associated morphological markings.

ERP research into agreement processing by native speakers of a language has shown some robust patterns. One of the most frequently reported associations across studies involving different languages is that between violations of morphosyntactic agreement and the emergence of the ERP component known as P600, a positive deflection peaking between 500 and 900 ms after stimulus onset, typically maximal over electrodes situated at left-parietal regions of the scalp (e.g., Alemán Bañón et al., 2012; Frenck-Mestre et al., 2008; Hagoort, 2003; Osterhout & Mobley, 1995).

This does not make the P600 a component specific to human syntax, but rather reflects sensitivity to some kind of rule-based structural knowledge, found both within and beyond language (e.g., Kuperberg, 2007; Patel et al., 1998; van de Meerendonk et al., 2013). Although violations of syntactic agreement have been shown to elicit a wider range of ERP patterns across populations (e.g., Friederici et al., 1996; Molinaro et al., 2015), which might reflect averaging effects over individual differences (Tanner, 2019), grand mean analyses in sufficiently large groups of native speakers almost invariably yield a P600 component. Nonnative speakers typically display more variability in their electrophysiological responses to agreement violations, which has been associated with differences in second language (L2) proficiency and typological distance between the first language (L1) and L2 (see Alemán Bañón et al., 2018, for an overview). At higher proficiencies, nonnative speakers who are familiar with a given syntactic agreement feature from their L1 (e.g., number, person, or gender agreement) have been found to display P600-like responses to agreement violations (e.g., Alemán Bañón et al., 2014, 2018; Foucart & Frenck-Mestre, 2012). By contrast, some novel (i.e., L2-only) features elicit P600 responses only in very high proficiency L2 speakers (Rossi et al., 2014).

Language Processing in Bilectal Speakers

Although there is some neurocognitive research on bilectal speakers, the existing literature has mainly examined phonetic/phonological and semantic processing (e.g., Bühler et al., 2017; Goslin et al., 2012; Lanwermeyer et al., 2016; Martin et al., 2016). A composite view from these studies suggests that (a) there is a difference in language processing between native and nonnative speakers of a dialect, and (b) the processing of phonetic/phonological and semantic features seems to depend on the amount of exposure to the given dialect.

Research investigating online processing of morphosyntactic dialectal variation is extremely scarce. To our knowledge, there are only two recent studies on the topic involving subregional dialects of American English. Zaharchuk et al. (2021) examined the processing of double modals (e.g., *might could*)—a construction that is often used in Southern United States English (SUSE)—by two groups of listeners: (a) those who are exposed regularly to SUSE (Southern) and (b) those who are not (non-Southern). In a comprehension task using ERPs, they found that both groups of listeners displayed a larger amplitude P600 to double modals than to single modal constructions (e.g., *could*). Findings thus suggest that the two types of listeners, regardless of their relative exposure to SUSE, processed double modals in a similar way, more in line with what might be expected for the standard (or unmarked) language variety. Such a finding leaves unclear, if not questionable, the extent to which individuals more exposed to SUSE actually have distinct representations for double modals, given the null effect. Garcia (2017) and Garcia et al. (2022) examined ERP responses to the omission of the third-person singular verb agreement marker -*s*, such omission being a prominent feature of African American Vernacular English. Monolectal speakers of Standard American English showed larger P600 effects when the verb agreement marker was omitted than bilectal speakers of African American Vernacular English and Standard American English. The offline behavioral results mirrored the ERP findings: Bilectal speakers accepted sentences with -*s* drop more often than monolectal speakers. As a result, these findings seem to indicate a distinct representation from that for Standard American English.

Taken together, the picture that emerges (admittedly from only two studies) is rather unclear. Although there are both methodological and sociolinguistic reasons for the differing results from the two studies (showing or not showing EEG evidence for distinct representations underlying the features), the fact that the studies were carried out in the United States (U.S.) could also be relevant. In the U.S. context, the sociolinguistic milieu differs greatly from other contexts of regional bilectalism found in different parts of the world. For example, the two subdialects of American English investigated in the studies clearly do not share equal status with Standard American English in the U.S., very differently from the case of Norwegian dialects in Norway. In the case of the U.S., it is often difficult to tease apart the effects of social stigma, prestige, identity, and salience attached to certain dialects from the effects of input, since more prestigious, standardized dialects tend to be more available in the listener's input. The dialectal landscape of Norway, as described in detail below, is an ideal setting to minimize this issue of dialectal status: No official spoken standard exists in Norway, and thus dialects are perceived as a part of one's identity and a sign of democracy and decentralization (Røyneland, 2009).

Norway as a Natural Laboratory

Varieties of Norwegian

Like most other countries in Western Europe, Norway must be characterized as multilingual. Currently, 18.5% of the population of 5.4 million are immigrants or children of two immigrants and speak a multitude of different languages (Statistics Norway, 2021). The main official language of the country is Norwegian, used in all sectors of life and throughout the country, and even this language has—for historical reasons—two mutually intelligible but distinct written varieties with unique grammatical systems: Bokmål and Nynorsk. Even though they enjoy equal judicial status at the national level, Bokmål is by far the most widespread variety across all sectors of public life: A modest 10–15% of the population are regular users of Nynorsk.

In addition to writing in the standard varieties, the practice of writing "in dialect" in private contexts in regional parts of Norway has become commonplace, especially in computer-mediated communication and social media (see Røyneland & Vangsnes, 2020, and references therein). In dialect writing, dialect users display and make use of local or regional phonological, morphosyntactic, and lexical forms (e.g., NN dialect writing displays dialect-specific number nonagreement on predicative adjectives, cf. (1–4) below). In other words, regional dialects in contemporary Norway are used in both spoken and written dimensions. This is important, as the stimuli in this study were written in NN dialect.

As for spoken Norwegian, there are four major dialect regions (Eastern, Western, Central ["Trøndersk"], and Northern Norwegian), which can be further subdivided into several subdialects. Quite uniquely, Norway has no official or even de facto spoken standard, although the Eastern Norwegian dialect spoken in and around the capital, Oslo, arguably has a particular status. This is witnessed, for instance, by the fact that this variety is spreading within Eastern Norway, causing dialect leveling and even virtual dialect shifts in many places.

Currently, regional or local dialects are spoken in both private and public contexts, including TV, media, universities, and parliament, irrespective of the audience. Thus, diglossic situations are rare in contemporary Norway as far as spoken language is concerned (Røyneland, 2009), and all Norwegians are, to a greater or lesser extent, exposed to the spectrum of Norwegian dialects.

However, there is an asymmetrical skewing of such exposure to Southern Eastern Norwegian due to the economic and political significance of the capital region and to the distribution of the population (the Oslo regional area is home to 1.7 million of the 5.4 million Norwegians). Consequently, Northern Norwegian dialect natives will have much more access and exposure to Southern Eastern dialect than the other way around.

Grammatical Variation in Norwegian

Although Norwegian spoken dialects are largely mutually intelligible, there are important grammatical differences that go well beyond vocabulary and accent. In some cases, the grammatical differences pertain to shared morphosyntactic categories where the dialects display narrow misalignment in terms of the specific configurations they license. In other words, something that is grammatical in one variety (e.g., agreement marking) can be markedly ungrammatical in the other (and vice versa). These differences constitute the perfect testing ground for our research questions.

One such case can be found when looking at NN number agreement in the nominal domain (in particular, noun–predicative adjective agreement) in contrast to most other varieties. This is illustrated below in (1) and (2), which show dialectally nonvarying masculine/neuter gender agreement (1a, b; 2a, b) and dialect-specific number (non) agreement (1c,2c) (def. = definite; m. = masculine; n. = neuter; sg. = singular; pl. = plural; Ø = null element; * indicates ungrammaticality). The data in (2) are written in so-called NN dialect writing (a nonstandardized but frequently used dialect orthography, commonly found in digital social communication, with which all participants were familiar). Regarding the number agreement patterns in (1c, 2c), most Norwegian dialects, as well as both written standards Bokmål and Nynorsk, have obligatory number agreement between a plural subject noun and a predicative adjective in copular constructions, as shown for Bokmål in (1c). However, in NN the adjective cannot be overtly inflected in the plural for the sentence to be grammatical (i.e., these forms have zero marking), as shown in (2c).

(1) *Bokmål*

a. Bil-en er fin/*fin-t.
Car-def.m.sg. is nice-m.sg./*n.sg.
"The car is nice."

b. Hus-et er fin-t/*fin.
House-def.n.sg. is nice-n.sg./*m.sg.
"The house is nice."

c. Bil-ene/Hus-ene er fin-e/*fin.
Car/House-def.pl. are nice-pl./*Ø
"The cars/houses are nice."

(2) *Northern Norwegian*

a. Bil-en e fin/*fin-t.
Car-def.m.sg. is nice-m.sg./*n.sg.
"The car is nice."

b. Hus-e e fin-t/*fin.
House-def.n.sg. is nice-n.sg./*m.sg.
"The house is nice."

c. Bil-an/Hus-an e fin/*fin-e.
Car/House-def.pl. are nice-Ø/*pl.
"The cars/houses are nice."

The variation in number agreement outlined in (1–2) above provides an ideal natural setting for studying bilectal processing. The NN dialect requires zero marking on adjectives in predicates, disallowing number agreement, but when writing in Bokmål (the common standard orthography), NN natives must apply number agreement, which is at odds with their native dialect. The opposite is potentially true of an individual who has lived in Northern Norway for a significant time, but moved there from another dialect region: Their native grammar (and their knowledge of the common written standardized varieties) requires number agreement at odds with the NN grammar to which they have been (amply) exposed in their current naturalistic context.

This is in sharp contrast to agreement properties that are the same across all dialects of Norwegian and standard written varieties, for example, gender as in (3) and (4) below. For gender, any given violation of agreement is equally ungrammatical in all dialects.

(3) *Bokmål*

Liv fortalte meg at hund-en hun trente var snill/*snil-t mot barn.
Liv told me that dog-def.m.sg. she trained was kind-m./*kind-n. to kids.

"Liv told me that the dog she trained was kind to kids."

(4) *Northern Norwegian*

Ho Liv fortalte mæ at hund-en ho trænte va snill/*snil-t mot unga.
Liv told me that dog-def.m.sg. she trained was kind-m./*weak-n. to kids.

"Liv told me that the dog she trained was kind to kids."

The present study uses written stimuli in NN like (4) to examine whether implicit measures of language processing—in this case, time-locked EEG activity—show that the number agreement induces a conflictive parse as opposed to the baseline condition of gender. This can provide crucial information about linguistic representation and processing in bilectals, whether natives of the target dialect tested or L2 bilectals of it.

The Present Study

The present study measures and regresses exposure to or experience with NN as a continuous variable in a diverse group of Norwegian speakers, ranging from born-and-raised Northern Norwegians to people who moved to Northern Norway from other dialectal regions later in life. We examine both offline behavior (grammaticality judgments) and implicit language processing (ERPs) in the NN dialect within two grammatical domains: (a) number agreement and (b) gender agreement. The research questions are formulated as follows:

i. How do Norwegians living in Northern Norway process gender and number agreement information in the NN written dialect mode?
ii. To what extent do exposure to NN and/or native speaker status in regard to NN modulate potential individual differences?
iii. What does answering (i) and (ii) mean for discussions of whether bilectalism is a proper subcase of bilingualism, and in particular how nonnative dialectal acquisition mirrors (or not) typical L2 bilingualism?

Recall that, whereas all dialects and written varieties show the same patterns for gender agreement, NN shows essentially the opposite pattern to (most) other dialects and the two written varieties (Bokmål and Nynorsk) for number agreement in the context we examine here. Absence of the agreement marker is obligatory (grammatical), and presence of the agreement marker is ungrammatical: for example, NN *bilan e fin/*fin-e* “the cars are nice-Ø/*pl.” At the behavioral level, we expect all subjects, irrespective of nativeness and/or degree of exposure to NN, to be able to provide judgments suggesting they have at least explicit knowledge of how the NN dialect works: where it accords with other varieties (gender) and where it differs from them (number). At the brain level, however, whereas we do not anticipate any NN nativeness- or exposure-related differences for gender, we do expect possible differences for number, which could go in several directions.

We expect everyone to show a P600 effect, because no matter which parsing strategy one uses (requiring or prohibiting number agreement, e.g., *bilan e fin/fin-e* “the cars are nice-Ø/pl.”), one of the presented forms is grammatical whereas the other is ungrammatical; the grammaticality is simply flipped across dialects (i.e., overt marking, which is consistent with most Norwegian dialects as well as the two written standards Bokmål and Nynorsk, or zero number marking, which is consistent with NN).

The relevant question is what triggers the P600 effect: Is it the overt marking (ungrammatical in NN) or the zero marking (ungrammatical in the other dialects of the non-NN natives)? In principle, it is possible that the driving force of the P600 is the same for all. This would mean that everyone (in our sample) uses the NN strategy to parse number grammaticality, where the signature is driven by the presence of overt agreement. This would indicate that all—natives and L2 bilectals alike—have (distinct) representations of NN and use them for processing NN input. Because we have a rather large spectrum of exposure to NN in our nonnative group, this would show that relatively little dialectal exposure is needed for one to process number marking in a NN-like manner. Alternatively, NN nativeness could predict P600 effects that are fully reversed between the two groups. For NN natives, the P600 effect would be elicited by instances of number agreement, ungrammatical in NN. For nonnatives, a P600 effect could be elicited by lack of agreement, ungrammatical in (relevant) non-NN dialects. If such a strict divide were to be evidenced, given that our nonnative group includes individuals with very high levels of exposure to NN in its native context, this would suggest that even significant exposure to nonnative dialects is inconsequential for developing distinct representations for it. In line with contemporary data showing a continuous effect of engagement with bilingual language use or exposure, we find the most likely outcome to be individual differences for the bilectals, evidencing that the NN grammatical contrast drives the P600 commensurable with increased exposure to and/or engagement with NN over time. If demonstrated, this would mean that nonnative bilectal acquisition is qualitatively similar to other instances of typical L2 acquisition.

Method

Participants

There were 111 native Norwegian participants tested for the current study. Since the NN pattern for the linguistic property under investigation is shared by certain Central dialects ("Trøndersk"), we excluded speakers born and/or raised in that dialect region. Since the property shows variation among certain Western Norwegian dialects (aligning in some cases with NN), we decided on inclusion for potential participants born and/or raised in this area based on a spoken sentence production task (see below) that revealed whether they omitted the plural marker in their native dialect in the relevant contexts. Six participants were excluded as a result, giving a final pool of 105 participants (mean age = 28.25 years, *SD* = 13.79, female = 74). Sixty-one participants were native speakers of NN who were born in Northern Norway (mostly in Troms County). Those who were not originally from Northern Norway (n = 44) and categorized as nonnative NNs were born in various parts of Norway (or, in the case of four participants, were born abroad but moved to other parts of Norway at the latest by age three).

Most participants received their primary education in the Bokmål written standard ($n = 101$), whereas the remaining four participants were educated in both Bokmål and Nynorsk. The average education level was 3.29, on a scale from 1 (*primary school*) to 5 (*postgraduate degree*; 13 scored 1 on the scale, 17 scored 2, 18 scored 3, 31 scored 4, and 26 scored 5). Participants had normal or corrected-to-normal vision, no reading disorders, and no history of neurological impairment.

Tasks

Language Social Background Questionnaire and Spoken Sentence Task

We adapted the Language Social Background Questionnaire (LSBQ; Anderson et al., 2018) to assess the participants' degree of use of and exposure to the NN dialect. The LSBQ is a questionnaire originally established to derive composite factor scores that represent overall levels of engagement with bilingualism in various domains. It includes a detailed description of bilingual usage patterns with different interlocutors across various settings and provides a continuous assessment of bilingual language use and engagement. Given that, to date, there is a lack of appropriate tools to assess engagement in bilectalism, instead of asking about exposure to one language versus the other (as indicated in the original version of the LSBQ), we adapted this questionnaire to reflect the participants' exposure to the NN dialect versus other dialects. A copy of this adapted questionnaire can be accessed via the following link: https://forms.gle/b4N8QJrnCQhiNaMq6. We then derived three individual factor scores via the Factor Score Calculator (see Administration and Scoring Manual: https://doi.org/10.6084/m9.figshare.3972486.v1): (a) the extent of NN dialect use at home, (b) NN dialect use in social settings, and (c) proficiency in the NN dialect. Finally, a composite score was computed by summing the factor scores weighted by each factor's variance, reflecting participants' overall engagement with the NN dialect. This composite score was used for further analyses in predicting language processing outcomes.

Assignment of native and nonnative NN speaker status for individuals was done based on the biographical and language history factors from the LSBQ in conjunction with a small spoken sentence production task. All participants who declared being born and raised in Northern Norway, at least to the age of six (applicable for a few who moved away from Northern Norway after the age of six only to return later in life), and who used the NN dialect in the sentence production task were assigned to the NN native group. This task, administered after completing the ERP experiment, asked participants to read out loud six sentences, written in Bokmål, in their native dialect (see Appendix S1 in the Supporting Information online). These sentences contained the critical structure—a predicative adjective agreeing with a plural subject noun—and were recorded in a separate sound file. For those who were categorized as NN nonnatives, the mean age of moving to Northern Norway was 21.41 years ($SD = 4.94$, range: 6 years to 37 years), and the average length of residence in Northern Norway was 8.92 years ($SD = 12.83$, range: a few months to 47 years).

Event-Related Potential Experiment

In the main task, participants read sentences in NN on a computer screen while electrophysiological activity was recorded at the scalp. As can be seen in the examples above, especially (3) and (4), the written sentences make it obvious that the modality is NN, due to both morphosyntactic and lexical differences between Bokmål and NN. Some of these cues are the use of a preproprial article in NN (*ho Liv* "she Liv"), specific lexis (e.g., Bokmål *barn* vs. NN *unga*, "children"), NN pronominal forms such as the oblique *mæ* "me" as opposed to *meg*, and inflectional differences such as noun definite–number marking (e.g., NN *bil-an* and Bokmål *bil-ene* "car"-def.pl). Crucially, relevant differences that characterize sentences as NN occur several times across all sentences, that is, whether or not the target domain (gender or number) is distinct in NN. Sentences were presented word by word, employing the rapid serial visual presentation method. The time from the onset of one word to the onset of the next was 450 ms, with no interstimulus interval. This was intended to enhance naturalness in reading pace and experience (see, e.g., Botella & Eriksen, 1992; Dambacher et al., 2012; Ditman et al., 2007) while still allowing for accurate time-locking of the ERPs.

Embedded into the EEG recording was a behavioral judgment task. After the last word of each sentence, participants were instructed to press the button we marked for "correct" (the left button) or "incorrect" (the right button) on the keyboard according to their judgment. The words "RIKTIG" ("correct") and "GALT" ("incorrect"), appearing on the same side of the screen as the relevant response button, were defined to the participants in the following way:

"RIKTIG" was appropriate if they believed a given sentence is grammatical in any one version or more of Norwegian and "GALT" if impossible in all versions of Norwegian. They were told that there was a time limit of 2.5 seconds to provide their response. Acceptance rate and reaction time were recorded as behavioral measures.

The experiment consisted of 300 sentences, of which 120 were experimental items and 180 were fillers. Experimental sentences belonged to one of four conditions, depending on the morphosyntactic property they probed and whether they were grammatical or ungrammatical in NN. We will refer to these conditions as gender agreement, gender nonagreement, number agreement, and number nonagreement. To avoid having the same participant see the agreement and mismatch versions of the same (gender or number) sentence, we constructed two experimental lists. A total of 60 sentences per condition were constructed and assigned to each list following a Latin square design, yielding 30 trials per condition per subject.

The structure of gender and number sentences was identical, as seen in Examples (5) and (6), respectively (see Appendix S2 in the Supporting Information online for a full list of target stimuli). Critical agreement was evaluated between the subject noun of the subordinate clause following the complementizer *at* and the predicative adjective appearing after the past-tense copula *va*. Violations in mismatch conditions were instantiated on the adjective and concerned only one feature (i.e., there were no double violations of gender and number). A final region of two to four words was added after the critical adjective to control potential wrap-up effects.

(5)	*Ho*	*Liv*	*fortalte*	*mæ*	*at*	***teksten***	*ho*	*skreiv*	*va*	***svak/*svakt***	*i*	*språket.*
	She	Liv	told	me	that	text-def.m.sg.	she	wrote	was	poor-m.sg./*poor-n.sg.	in	language.

"Liv told me that the text she wrote was not well written."

(6)	*Ho*	*Liv*	*fortalte*	*mæ*	*at*	***gjestan***	*ho*	*hadde*	*va*	***glad/*glade***	*i*	*mat.*
	She	Liv	told	me	that	guest-def.m.pl.	she	had	were	happy-∅/*happy-m.pl.	in	food.

"Liv told me that the guests she had were fond of food."

Gender violations were manipulated between masculine and neuter to obtain robust effects, given that the feminine gender is currently in the process of disappearing in several dialects of Norwegian (e.g., Busterud et al., 2019; Lohndal & Westergaard, 2021; Rodina & Westergaard, 2021). The number of masculine and neuter nouns was counterbalanced. All subject nouns in the number conditions were plural, since this is the only context in which NN differs from other dialects for this property. Number nonagreement sentences contained a zero-marked predicative adjective (grammatical in NN), whereas number agreement sentences displayed an overtly marked plural adjective (ungrammatical in NN, grammatical in other dialects). ERPs were time-locked to the onset of the critical word (the adjective).

Fillers had a similar structure up to the complementizer. This was followed by a noun phrase consisting of an attributive adjective and a noun, preceded in half the sentences by the determiner *det/den*. Only a sixth of the filler sentences were patently ungrammatical, because the adjectives employed do not tolerate article omission. The rest of the filler sentences ($n = 150$) were therefore on a scale of acceptability. These materials were designed to investigate the effect of determiner omission as a function of adjective type in Norwegian double definite constructions.

Procedure

Upon arrival, participants read and signed an informed consent form. After completing the LSBQ, they were escorted into a sound-attenuated glass booth where the ERP experiment took place. After EEG cap preparation and system setup, the experiment started. The 300 sentences were divided into 10 blocks of 30 sentences. Participants were encouraged to take a rest at every break and continue at their own pace. Sentences were pseudorandomized so that each block contained three sentences from each condition (the current experiment had four conditions; fillers had six), and two sentences from the same condition could be, at the closest, 10 items apart.

EEG Recording and Processing

The EEG signal was continuously recorded from the scalp using 32 active electrodes (ActiCap, Brain Products, Inc.) fitted in an elastic cap organized according to the international 10–20 system. AFz served as the ground electrode. EEG recording was referenced online to electrode FCz and re-referenced offline to the average mastoids (TP9/10). The frontoparietal electrodes FP1 and FP2, located above the eyebrows, were used to monitor eye blinks. Impedance was kept below 20 kΩs for all electrodes. The recordings were amplified by a LiveAmp amplifier (Brain Products, Inc.) with a bandpass filter of 0.01–200 Hz and digitized continuously at a sampling rate of 500 Hz. Preprocessing of the EEG data was performed on Brain Vision Analyzer 2.0 (Brain Products, Inc.). All trials, regardless of accuracy in the grammaticality judgment task, were considered for analysis (following procedures from Mickan & Lemhöfer, 2020; Tanner et al., 2013).

Offline data were filtered with a bandpass filter of 0.1–30 Hz. The continuous EEG was then segmented into 1,500-ms epochs with reference to the critical word (adjective). Epochs contained a 300-ms prestimulus baseline and ended 1,200 ms after stimulus onset. Trials were manually inspected for artifacts (drifts, excessive muscle artifact, blinks, blocking, etc.). Rejecting trials with artifacts resulted in the exclusion of 3.23% of trials. The average number of trials kept did not differ by condition (number grammatical: 24.5; number ungrammatical: 24.4; gender grammatical: 24.5; gender ungrammatical: 24.3; all ps $> .1$). The remaining epochs were baseline-corrected relative to the 300-ms prestimulus baseline and averaged by condition for each participant. Due to technical issues, for 33 participants, roughly 31% of the aggregate, only 50% of the trials (equally across all conditions) were input. Given the excellent inclusion rate overall after preprocessing, this still yielded an average across all participants of 81.75% of trials for the grand averaging, which is a robust result.[2] On the basis of the literature on the processing of nonlocal morphosyntactic agreement, ERPs were identified by measuring mean amplitudes in the 500–900-ms time window, where late positive components elicited by morphosyntactic processing, such as the P600, tend to be found for both native and nonnative speakers (e.g., Alemán Bañón et al., 2012, 2014; Kaan, 2007). Analyses were conducted on a subset of the electrodes based on seven regions of interest (left anterior: F3/7, FC1/5; right anterior: F4/8, FC2/6; left posterior: CP1/5, P3/7; right posterior: CP2/6, P4/8; midline anterior: Fz; midline medial: Cz; midline posterior: Pz, Oz; see, e.g., Tanner, 2019, for a similar setup).

Results

Behavioral Data

We first report the results from the behavioral data. The descriptive statistics of the proportion of acceptance and reaction time collapsed over condition (gender vs. number) and agreement (agreement vs. nonagreement) are illustrated in Figure 1. It is important to emphasize here that, for the gender condition, "nonagreement" refers to neuter-to-common (masculine and feminine, which are collapsed in dialects with two-gender systems) violations that are shared across both NN and other Norwegian dialects. However, for the number condition, "nonagreement" trials refer to nonagreement between predicative adjective and noun in number, which is grammatical in NN but ungrammatical in other dialects.

We report statistical analyses on the proportion of acceptance only, since we are interested in the participants' grammatical judgments themselves rather than how long it took them to arrive at those judgments.[3] We ran a generalized linear mixed-effects model with binary choice (correct or incorrect)

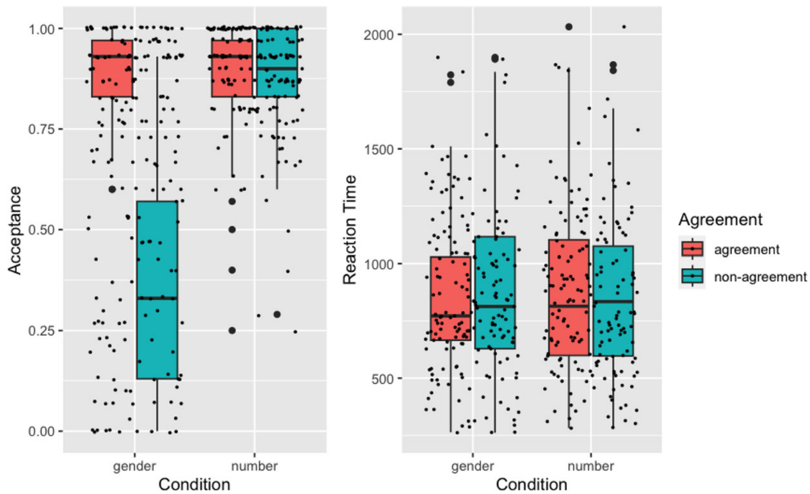

Figure 1 Descriptive illustration of the proportion of acceptance and reaction time from the behavioral data using boxplot.

as a dependent variable. Condition (gender, number), agreement (agreement or nonagreement), LSBQ composite score, nativeness (native NN, nonnative), and their four-way interactions were included as fixed effects, as well as subject as a random intercept and agreement and condition as by-subject slopes. We used the sjPlot package in R (Lüdecke, 2022) to generate diagnostic plots and confirm model assumptions. We used the mixed function in the afex package in R (Singmann et al., 2015) to run a likelihood ratio test. The categorical variables were sum coded, and the LSBQ composite score was centered around the mean. Pairwise post hoc comparisons with Tukey contrasts were conducted using the emmeans package in R (Lenth et al., 2019). The alpha that we will use for interpreting statistical significance throughout the manuscript is .05.

The output of the likelihood ratio tests (Table 1) shows that there was a significant main effect of agreement ($p < .001$) and condition ($p < .001$) and a significant interaction between agreement and condition ($p < .001$). That is, the difference in the rate of acceptance between agreement and nonagreement trials was larger for gender than number. However, this effect was not modulated by (a) the participants' exposure to NN (LSBQ composite score; $p = .05$) or (b) whether they were NN natives or nonnatives (nativeness; $p = .19$) or (c) the interaction of both ($p = .14$). The summary of the model with odds

Table 1 Likelihood ratio test results from generalized linear mixed-effects model of behavioral (acceptance) data

Effect	*LR*	*p*
Agreement	60.90	**<.001**
Condition	50.83	**<.001**
Composite_c	0.52	.47
Nativeness	0.00	>.99
Agreement Condition	114.83	**<.001**
Agreement × Composite_c	0.00	.94
Condition × Composite_c	0.18	.67
Agreement × Nativeness	0.15	.70
Condition × Nativeness	0.43	.51
Composite_c × Nativeness	0.39	.53
Agreement × Condition × Composite_c	3.69	.05
Agreement × Condition × Nativeness	1.67	.19
Agreement × Composite_c × Nativeness	1.99	.15
Condition × Composite_c × Nativeness	1.29	.25
Agreement × Condition × Composite_c × Nativeness	2.16	.14

Note. Model: Acceptance ∼ Agreement × Condition × Composite_c × Nativeness + (Agreement + Condition|Subject). Composite_c = Language Social Background Questionnaire composite score. Estimates in boldface are statistically significant at $\alpha = .05$. *LR* refers to likelihood ratio statistics.

ratios is included in Table S3.1 in Appendix S3 in the Supporting Information online.

Event-Related Potential Data

For the ERP analysis, we examined the modulatory effects of exposure to NN and native speaker status on ERP mean amplitudes at the time window of interest (500–900 ms). We built a linear mixed-effects model with amplitude as the dependent variable and the following as fixed effects: condition (gender, number), agreement (agreement, nonagreement), LSBQ composite score, nativeness (native NN, nonnative), and scalp region of interest (left-anterior, right-anterior, left-posterior, right-posterior, midline-anterior, midline-central, and midline posterior), as well as a four-way interaction between condition, agreement, LSBQ composite score, and nativeness. The random structure of these models contained random intercepts for subject and electrode and by-subject slopes for condition and agreement. We used the sjPlot package (Lüdecke, 2022) to generate diagnostic plots and confirm model assumptions.

Table 2 Likelihood ratio test results from linear mixed-effects model of event-related potential data

Effect	*LR*	*p*
Agreement	4.36	.03
Condition	12.86	**<.001**
Composite_c	1.36	.24
Nativeness	2.31	.12
ROI	32.72	**<.001**
Agreement × Condition	209.82	**<.001**
Agreement × Composite_c	0.07	.79
Condition × Composite_c	1.27	.26
Agreement × Nativeness	1.10	.29
Condition × Nativeness	0.59	.44
Composite_c × Nativeness	2.53	.11
Agreement × Condition × Composite_c	0.71	.40
Agreement × Condition × Nativeness	0.01	.90
Agreement × Composite_c × Nativeness	0.42	.51
Condition × Composite_c × Nativeness	2.57	.10
Agreement × Condition × Composite_c × Nativeness	37.18	**<.001**

Note. Model: Amplitude ∼ Agreement × Condition × Composite_c × Nativeness + ROI + (Agreement + Condition|Subject). Composite_c = Language Social Background Questionnaire composite score; ROI = scalp region of interest. Estimates in boldface are statistically significant at $\alpha = .05$. *LR* refers to likelihood ratio statistics.

Upon visual inspection, we removed outliers with a standardized residual at a distance greater than 2.5 standard deviations from 0, using the romr.fnc function from the LMERConvenienceFunctions package in R (Tremblay, 2020). In total, 111 data points were removed. As in the behavioral data analysis, we used the mixed function in the afex package (Singmann et al., 2015) to run a likelihood ratio test. The categorical variables were sum coded, and the LSBQ composite score was centered around the mean. Pairwise post hoc comparisons with Tukey contrasts were conducted using the emmeans package (Lenth et al., 2019). The summary of this model with *b* coefficients is included in Table S3.2 in Appendix S3 in the Supporting Information online.

As shown in Table 2, there was a significant interaction between condition and agreement, and the post hoc comparisons revealed that nonagreement trials elicited a larger amplitude overall for gender than agreement trials (agreement − nonagreement: $b = -1.45$, $p < .001$), whereas no differences between agreement and nonagreement trials were found for number (agree-

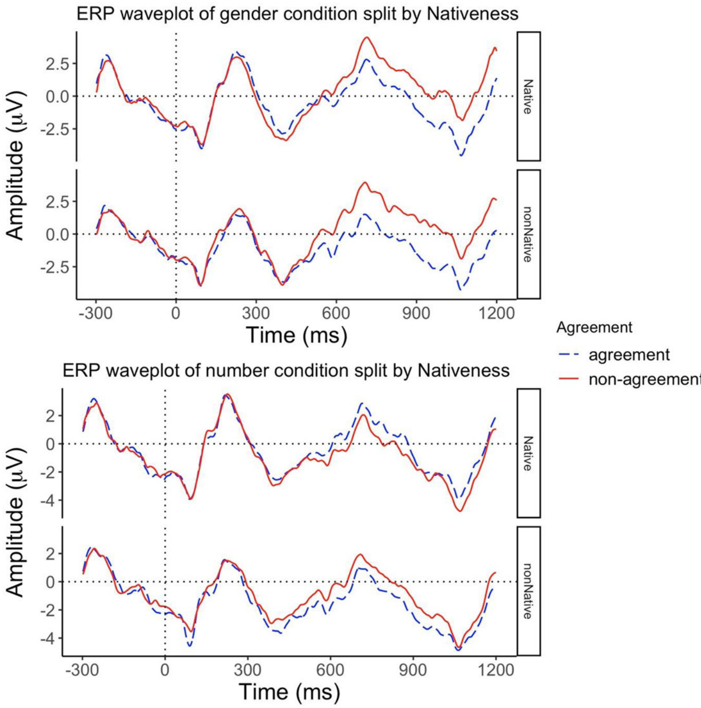

Figure 2 ERP waveforms recorded at electrode Pz for gender and number from native and non-native Northern Norwegian participants reading agreement sentences (blue dotted) and non-agreement sentences (red solid) in Northern Norwegian.

ment − nonagreement: $b = .43$, $p = .07$), because native and nonnative grammars with contrasting number agreement patterns cancel each other out at the aggregate level. This is shown in Figure 2, where the groups are split up by native speaker status. Here we can see a different pattern in the ERP waveforms between gender and number. Both native NNs and nonnative NNs show greater amplitude for nonagreement trials than agreement trials for gender. However, native NNs and nonnative NNs demonstrate a flipped P600 effect for number: That is, whereas native NNs show greater amplitude for number agreement than for nonagreement (ungrammatical in NN but grammatical in other dialects), nonnative NNs elicit greater amplitude for number nonagreement than for agreement (grammatical in NN but ungrammatical in other dialects).

The topographical maps in Figure 3, which plot the distribution of the agreement effect in gender (Panel A) and number (Panel B) at the aggregate level (not controlling for nativeness) provide visual confirmation: a broadly

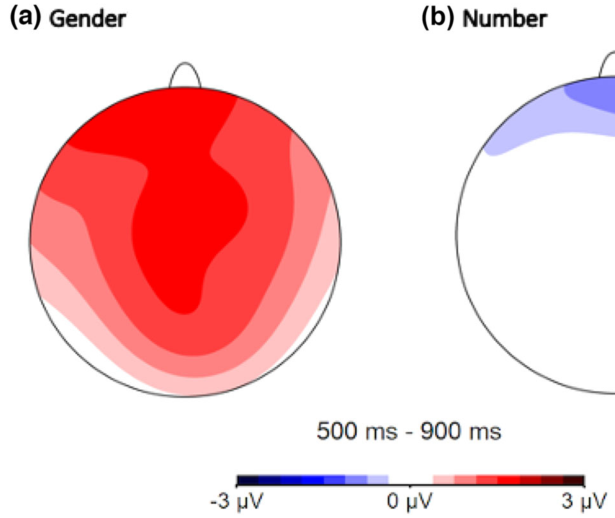

Figure 3 Topographical maps of A. the Gender non-agreement – Gender agreement difference, B. the Number non-agreement – Number agreement difference for the 500–900 ms time window.

distributed positivity for gender, and neutral polarity across most electrode sites for number, whereby the positivity for number agreement in the native NN and positivity for number nonagreement in the nonnative NN has a cancelling-out effect.

As demonstrated in Figures 2–3, we now know that nativeness plays a meaningful (but not deterministic) role in modulating number agreement ERPs, but we have not yet considered how prolonged exposure to NN in nonnatives influences the brain responses to NN-specific grammatical patterns. The significant four-way interaction between condition, agreement, LSBQ composite score, and nativeness further illustrates that LSBQ composite score, representing exposure to and engagement with the NN dialect, indeed significantly influences differences in ERP responses between native and nonnative groups.

As illustrated by Figure 4, Panel A, both NN nonnatives and NN natives show a clear P600 effect with greater amplitude for gender nonagreement than for agreement, and this effect is maintained across the spectrum of exposure to NN. However, in the number condition (Figure 4, Panel B), NN nonnatives and NN natives show a different pattern. That is, native speakers of NN, regardless of their exposure to NN, show a P600 effect with greater amplitude

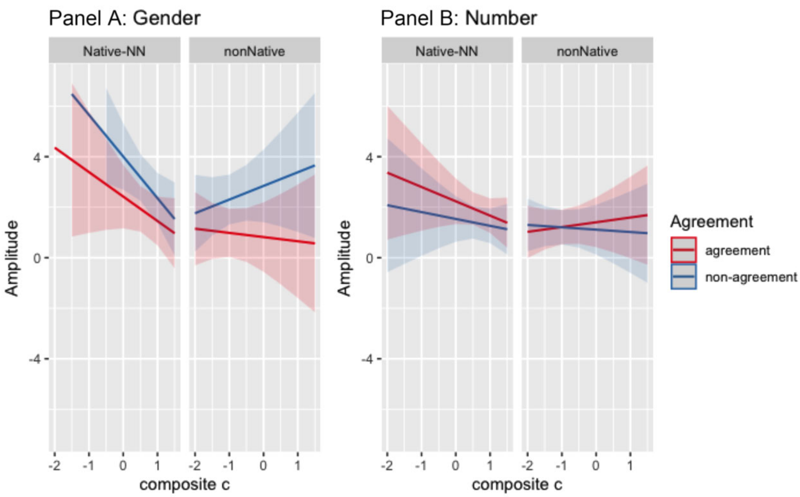

Figure 4 Illustration of the four-way interaction between Condition (gender, number), Agreement (agreement, non-agreement), Composite score (indicated as composite_c), and Nativeness (nonNative vs. Native-NN) on P600 component in the ERP data. Panel A shows the effect of Composite score and Nativeness on P600 effect for gender, and Panel B shows the effect of Composite score and Nativeness on P600 effect for number. The shaded areas represent 95% Confidence Intervals. Agreement is indicated in red and nonagreement in blue.

for number agreement, which is ungrammatical in their dialect than for nonagreement. In contrast, nonnative NNs appear to process number differently depending on their experience with NN, whereby a positive effect (with greater amplitude for number agreement) becomes more pronounced toward participants with the most NN exposure, whereas at the other end (least exposure to NN), it is nonagreement sentences that elicit more positive responses. In order to break down this significant four-way interaction, we ran a three-way interaction between agreement, LSBQ composite score, and nativeness, splitting the data by condition (gender and number). We found a significant interaction between agreement, LSBQ composite score, and nativeness for the number model ($LR = 8.64$, $p = .003$), but, as expected, not for the gender model ($LR = 2.31$, $p = .12$; see Table S3.3 and Table S3.4 in Appendix S3 in the Supporting Information online). These results suggest that being a NN native and having more exposure to the NN dialect matter for processing number agreement in NN (since NN and non-NN dialectal grammars treat number agreement differently) but not for gender agreement processing (since all Norwegian dialects equally require masculine/neuter gender agreement).

Discussion

In an effort to understand how bilectalism might impact grammaticality judgments, online processing, and their correspondence, the present study investigated the processing of gender and number agreement in a visual sentence processing ERP experiment conducted in the NN dialect. Our participants included native speakers of the NN dialect and speakers of other Norwegian dialects who moved to Northern Norway later in life. Given the specifics of the general Norwegian context, where the written standards (Bokmål and Nynorsk) are imparted to all via education, yet the colloquial dialect of the area where someone is raised will be their default spoken language, it is fair to say that all our participants are bilectal, if not multilectal. As a result, we did not take for granted a priori that anyone, even the NN natives, would show clear judgments and/or ERP signatures of ungrammaticality in the domains of interest for stimuli presented in NN, given that knowledge of Bokmål and/or Nynorsk can conflict with NN. Recall that the critical manipulation of our study involved the juxtaposition of two grammatical structures that vary in configurational overlap. Since gender agreement patterns (yielding neuter–masculine gender violations) are shared across all dialects, we predicted that everyone would offer the expected grammaticality judgments and classic ERP signatures (P600 responses): effects driven by nonagreement. Both judgment and ERP data confirmed these predictions straightforwardly.

Number agreement between a subject noun and a predicative adjective, however, differs between NN and the written standards and other spoken dialects. Therefore, we first sought to check whether our participants were able to correctly judge the opposite pattern of NN (obligatory lack of agreement, which is patently ungrammatical in most other Norwegian dialects), and, indeed, how they processed relevant sentences online. Behaviorally, unlike the case of gender, where all participants showed a distinction between grammatical and ungrammatical agreement, none of the participants showed such a distinction for number. This was likely a by-product of the fact that we asked whether this was a sentence that could be produced by a native Norwegian. We asked the question in this way to avoid any prescribed judgments on "correctness." After all, it could have been the case that some participants, especially nonnatives of NN, would take lack of agreement to be ungrammatical due to opinion or preference, even if they knew that it was the norm in

NN. Given that the sentences were clearly provided in NN, we expected the judgments to reflect the NN rules, at least for the subgroup of NN natives, which would mean roughly the same differences in judging gender violations as number ones in accord with NN. This was not, however, what we found. Across the board, whereas all participants showed the expected pattern for gender, they equally accepted sentences with and without number agreement. This was true irrespective of being a NN native and/or of relative exposure to NN.

This finding is interesting because it indicates that (a) all speakers are accepting of lack of agreement where such agreement is obligatory in non-NN dialects, yet (b) all still accept agreement even though the sentences are clearly in NN. The most likely explanation, given that agreement in this context is rather salient, is that our participants answered our question in a literal sense: Given at least explicit knowledge that some dialects of Norwegian allow for agreement and others allow for nonagreement, they could imagine both sentences being uttered by some Norwegian natives. What we can discern from this, regardless, is that even the nonnative NN participants know that lack of number agreement is possible: knowledge which could only be obtained via exposure to NN. In the future, checking this by running the same experiment with people who have not had any significant exposure to NN, for example, in Oslo, would put this assertion to a more rigorous test.

Returning to Figure 4, we can see that the picture of the online processing results is also interesting and yet does not necessarily mirror the judgment data just discussed. This was in some ways predictable given the very nature of how ERP experiments like this must be set up—in a comparative difference paradigm—and the way in which pattens of grammaticality vary between NN and other Norwegian dialects. After all, the relevant comparison of sentences with and without number agreement should always yield a P600 effect no matter which Norwegian grammar a participant is using to parse the sentences, precisely because one of the sentence types is ungrammatical in all the grammars—just not the same contrast across the board. Because the ungrammatical sentence type varies by dialect and our goal was to assess whether the NN grammar is the driving force, we should expect the P600 (amplitude) to be driven only by number agreement if everyone were using a NN grammar for parsing, because this is the ungrammatical form in NN. Alternatively, the effect should be driven by lack of agreement if everyone were using a non-NN grammar, because this is ungrammatical in the other dialects. From what we can see, however, for Figure 4, Panel B, and unlike what we saw for gender in Figure 4, Panel A, it is clearly not the case that everyone is using the same grammatical configuration to parse these sentences, despite performing in the same way on judgments.

The P600 effect is reversed in polarity across the spectrum of experience with NN. As one might expect of an individual using NN to parse these sets of sentences, a positive effect becomes more pronounced toward the most NN extreme of the scale, whereas on the other end (as a function of less exposure to NN), it is nonagreement sentences that elicit more positive responses. Perhaps unsurprisingly, for NN natives (see Figure 4, Panel B left) agreement elicits higher amplitude than nonagreement regardless of their exposure to NN. This tells us that, despite being fully competent in the written standard(s), any potential interference, at least at the aggregate level, does not overwhelm the processing of their native dialect. This also provides some evidence for what we argued above with respect to the behavior judgments. When providing a judgment, one can offer a more "considered" response to the way we asked the question (i.e., accept both number agreement and nonagreement), but the real-time patterns of the NN natives align with the claims that NN, in fact, disallows number agreement in the relevant context. This, however, is not true of the NN nonnatives (see also Figure 4, Panel B right). It is in this subgroup of participants that we can really appreciate the effect of NN exposure. With increasing experience with NN, the NN distinction begins to drive the P600 effect in a qualitatively similar fashion as in the native NN cohort. This pattern tells us that acquisition of a new dialectal grammar is possible, and that exposure matters for this, as it does for typical adult L2 bilingualism. Moreover, we should note that this interaction is mainly deriving from the slope of the agreement trials: NN nonnatives elicit higher amplitude for agreement trials with increasing NN exposure, whereas the amplitude of nonagreement trials remains stable across the range of exposure with NN. This is not surprising, given that this experiment was conducted in the (written) NN dialect mode, and thus, agreement trials indeed involved the ungrammatical structure. Thus, with more exposure to NN, NN nonnatives are more likely to become sensitive to ungrammaticality in their nonnative dialect.

As can be seen in Figure 4, however, it is not the case that for the nonnative NN participants we have two discernible groupings: Although less and more exposure to NN drives the picture of reversed polarity, it is also clear that the lines cross toward the middle of the scale. There are at least two possible accounts for this.

First, it might suggest that participants with only moderate experience with NN somehow did not process agreement and nonagreement sentences significantly differently. Although this is in line with the overall judgment data, we believe such an explanation to be unlikely, and not well motivated. After all, these are native speakers of Norwegian, and the NN dialect is mutually intelligible with their other dialects—which explains why no one, including those with the least amount of experience with NN, reported any problems in comprehension. For those participants at the lower end of the experience scale, a difference between agreement and nonagreement is patent in the ERP data. Hence, why would an individual with more exposure—though perhaps not enough to show the native NN pattern—stop processing these distinctions? This would be completely unnatural, given that there is no dialect of Norwegian that does not grammaticalize either obligatory agreement or obligatory lack of agreement.

A second possibility is that the observed pattern suggests an intermediary stage in bilectal acquisition whereby acquisition of the new target, in this case NN, does not preempt influence from its competitors when the learner is processing the target. In principle, this is no different from what has been documented in the adult L2 acquisition literature and labeled as the preemption effect (Kush & Dahl, 2022; Rothman & Iverson, 2013; Trahey & White, 1993). This is a very likely explanation considering the data in Figure 4, which show that some nonnative NN learners indeed have the native pattern, but that the likelihood of this correlates to being an individual in the upper echelons of exposure to NN. If this explanation is on the right track, then, those with moderate NN exposure appear to not process the relevant distinction because they are accessing NN for exemplars with a lack of number agreement and their other grammar(s) for exemplars appearing with agreement.

Our results are in line with those of Lundquist and Vangsnes (2018) in another Norwegian bilectalism study, which examined whether Sogn–Oslo bilectals with low and high exposure to and alignment with the Sogn dialect can use gender information in both dialect modes predictively. There is an obligatory three-gender (Masculine, Feminine, Neuter) split in the Sogn dialect whereas the Oslo dialect only makes a two-gender distinction (Common, Neuter). They found that the group of participants who were more traditional speakers of the Sogn dialect (as evidenced by self-reports and production tasks) reliably used all gender cues when a processing task was presented in the Sogn dialect, whereas the less traditional group did not exploit the masculine–feminine distinction. When presented with the Oslo dialect, neither group exploited the masculine–feminine distinction. Such a modulatory effect of dialect alignment on grammatical processing is also corroborated in our findings. Our study presses this further, going beyond the categorical distinction of high exposure versus low exposure by operationalizing dialectal experience as a gradient, continuous variable. Only by doing so were we able to see how the interaction effect unfolded. It was *not* the case that all participants showed the expected P600 effect (i.e., higher amplitude for agreement trials than for nonagreement) and that the magnitude of this effect was modulated by exposure (i.e., that there were quantitative differences). Instead, we witnessed an inverse relationship between exposure and number (mis)match (i.e., qualitative differences).

This indicates that there is some type of thresholding effect for bilectalism as there is for adult L2 acquisition. The composite LSBQ score we used, which has often been taken as an overall proficiency score in relevant research (Grundy et al., 2017; Hermanto et al., 2012), seems to translate to either (a) where along the developmental continuum in the acquisition of the target the learner is or (b) what the bilectal grammar looks like at an ultimate attainment point under various degrees of exposure or engagements conditions. This is also what we would expect to see in more typical L2 acquisition. Thus, we can say that bilectalism is similar in many ways to traditional L2 acquisition.

Limitations and Future Directions

The grammatical processing patterns we see with bilectals in Northern Norway mirror those of L2 learners. However, our examination was confined to one dialect mode. Based on these results, we do not know whether individual bilectals in this study would display distinct processing profiles if tested in other Norwegian varieties to which they are sufficiently exposed. To what degree do Norwegian bilectals (like bilinguals) show evidence of distinct grammatical representations for other Norwegian dialects and/or for other Norwegian written varieties they are engaged with (e.g., Norway's two official written standards, Bokmål and Nynorsk)? To answer this question, future research can test bilectals in two dialect modes, for instance, contrasting NN dialect writing and Bokmål, which display diametrically opposed grammaticality for the number agreement distinction we manipulated in this study (i.e., obligatory number agreement in Bokmål vs. prohibited number agreement in NN). Testing in two dialect modes would allow us to discern the extent to which bilectals adjust their processing strategies depending on the dialect input. It can also reveal how potential processing differences are further influenced by individual-level variation in bilectal exposure, akin to the variations identified in the present study.

The present work serves as a reminder of an important point: Focusing on binary oppositions and group comparisons is limiting. Doing so may conceal the fact that main effects often emerge from a variety of bilingual behaviors within one group or, vice versa, that critical group-internal variation can be obscured and neutralized at the aggregate level. This is an important point that applies well beyond the confines of bilectalism, that is, it applies regardless of the type of "bilingual" in focus. Failing to consider individual level of engagements with a language or a dialect, one can be misled to presume that different groups of speakers represent the same monolithic underlying population. This, in turn, can lead to erroneous overgeneralizations that (inadvertently) ignore systematicities in how speakers from distinct types of bilingualism can overlap or differ at the individual level (e.g., Beatty-Martínez & Titone, 2021; DeLuca et al., 2019; Leivada et al., 2021; Navarro-Torres et al., 2021).

Conclusion

In sum, the processing of NN dialect seems to differ depending on exposure to it: one does not need to be a native of NN to process its specific grammatical properties in a native-like manner. As we have discussed in some detail, we interpret the present data as supporting the view that bilectalism can be considered a proper subcase of bilingualism given how it resembles patterns reminiscent of typical L2 bilingualism. Like L2 acquisition, the present study demonstrates that bilectalism exists on a continuum whereby grammatical processing is modulated by an interaction of converging factors related to linguistic experience. Our findings also present evidence that merely being able to comprehend input or being able to judge sentences as (in)correct from closely related dialects does not *a priori* make an individual bilectal.

Final revised version accepted 22 June 2023

Open Research Badges

This article has earned Open Data and Open Materials badges for making publicly available the digitally-shareable data and the components of the research methods needed to reproduce the reported procedure and results. All data and materials that the authors have used and have the right to share are available at https://github.com/MakiKubota/Nordnorsk.git and as Supporting Information online. All proprietary materials have been precisely identified in the manuscript.

Notes

1 In using "native" or "nativeness" throughout the article, we are simply referring to the language(s) acquired on the basis of naturalistic exposure in early childhood.

2 When we removed these participants with only 50% of the data and ran the analyses, we still found a significant four-way interaction between agreement (agreement, nonagreement), condition (gender, number), composite score, and nativeness ($p = .045$), indicating that the composite score modulates the P600 effect for number and gender differently, and this pattern differs between NN natives and nonnatives. This interaction effect became more robust when we included the participants with only 50% of the data ($p < .001$). For analyses, see the R script, which can be accessed via https://github.com/MakiKubota/Nordnorsk.git

3 We provide the analysis of the reaction time data in the R script, which can be accessed via https://github.com/MakiKubota/Nordnorsk.git. In sum, there was no significant interaction between agreement and condition ($p = .70$), indicating that there is no difference in P600 effect between processing of gender and number. Moreover, the differences in reaction time between agreement and nonagreement sentences for gender and number are not modulated by exposure (i.e., composite score) or by nativeness.

References

Alemán Bañón, J., Fiorentino, R., & Gabriele, A. (2012). The processing of number and gender agreement in Spanish: An event-related potential investigation of the effects of structural distance. *Brain Research*, *1456*, 49–63. https://doi.org/10.1016/j.brainres.2012.03.057

Alemán Bañón, J., Fiorentino, R., & Gabriele, A. (2014). Morphosyntactic processing in advanced second language (L2) learners: An event-related potential investigation of the effects of L1–L2 similarity and structural distance. *Second Language Research*, *30*(3), 275–306. https://doi.org/10.1177/0267658313515671

Alemán Bañón, J., Fiorentino, R., & Gabriele, A. (2018). Using event-related potentials to track morphosyntactic development in second language learners: The processing of number and gender agreement in Spanish. *PLOS ONE*, *13*(7), Article e0200791. https://doi.org/10.1371/journal.pone.0200791

Anderson, J. A., Mak, L., Chahi, A. K., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, *50*(1), 250–263. https://doi.org/10.3758/s13428-017-0867-9

Beatty-Martínez, A. L., & Titone, D. A. (2021). The quest for signals in noise: Leveraging experiential variation to identify bilingual phenotypes. *Languages*, *6*(4), Article 168. https://doi.org/10.3390/languages6040168

Botella, J., & Eriksen, C. W. (1992). Filtering versus parallel processing in RSVP tasks. *Perception & Psychophysics*, *51*(4), 334–343. https://doi.org/10.3758/bf03211627

Bühler, J. C., Waßmann, F., Buser, D., Zumberi, F., & Maurer, U. (2017). Neural processes associated with vocabulary and vowel-length differences in a dialect: An ERP study in pre-literate children. *Brain Topography*, *30*(5), 610–628. https://doi.org/10.1007/s10548-017-0562-2

Busterud, G., Lohndal, T., Rodina, Y., & Westergaard, M. (2019). The loss of feminine gender in Norwegian: A dialect comparison. *Journal of Comparative Germanic Linguistics*, *22*, 141–167. https://doi.org/10.1007/s10828-019-09108-7

Chambers, J. K., & Trudgill, P. (1998). *Dialectology*. Cambridge University Press. https://doi.org/10.1017/CBO9780511805103

Dambacher, M., Dimigen, O., Braun, M., Wille, K., Jacobs, A. M., & Kliegl, R. (2012). Stimulus onset asynchrony and the timeline of word recognition: Event-related potentials during sentence reading. *Neuropsychologia*, *50*(8), 1852–1870. https://doi.org/10.1016/j.neuropsychologia.2012.04.011

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, *116*(15), 7565–7574. https://doi.org/10.1073/pnas.1811513116

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology*, *44*(6), 927–935. https://doi.org/10.1111/j.1469-8986.2007.00593.x

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, *66*(1), 226–248. https://doi.org/10.1016/j.jml.2012.02.009

Frenck-Mestre, C., Osterhout, L., McLaughlin, J., & Foucart, A. (2008). The effect of phonological realization of inflectional morphology on verbal agreement in French: Evidence from ERPs. *Acta Psychologica*, *128*(3), 528–536. https://doi.org/10.1016/j.actpsy.2007.12.007

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1219–1248. https://doi.org/10.1037/0278-7393.22.5.1219

Garcia, F. M. (2017). *Brain responses to contrastive and noncontrastive morphosyntactic structures in African American English and Mainstream American English: ERP evidence for the neural indices of dialect*. Unpublished dissertation. Columbia University.

Garcia, F. M., Shen, G., Avery, T., Green, H. L., Godoy, P., Khamis, R., & Froud, K. (2022). Bidialectal and monodialectal differences in morphosyntactic processing of AAE and MAE: Evidence from ERPs and acceptability judgments. *Journal of*

Communication Disorders, *100*, Article 106267. https://doi.org/10.1016/j.jcomdis.2022.106267
Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain and Language*, *122*(2), 92–102. https://doi.org/10.1016/j.bandl.2012.04.017
Grundy, J. G., Chung-Fat-Yim, A., Friesen, D. C., Mak, L., & Bialystok, E. (2017). Sequential congruency effects reveal differences in disengagement of attention for monolingual and bilingual young adults. *Cognition*, *163*, 42–55. https://doi.org/10.1016/j.cognition.2017.02.010
Hagoort P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, *15*(6): 883–899. https://doi.org/10.1162/089892903322370807
Hermanto, N., Moreno, S., & Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: How bilingual? *International Journal of Bilingual Education and Bilingualism*, *15*(2), 131–145. https://doi.org/10.1080/13670050.2011.652591
Kaan, E. (2007). Event-related potentials and language processing: A brief overview. *Language and Linguistics Compass*, *1*(6), 571–591. https://doi.org/10.1111/j.1749-818X.2007.00037.x
Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49. https://doi.org/10.1016/j.brainres.2006.12.063
Kush, D., & Dahl, A. (2022). L2 transfer of L1 island-insensitivity: The case of Norwegian. *Second Language Research*, *38*(2), 315–346. https://doi.org/10.1177/0267658320956704
Lanwermeyer, M., Henrich, K., Rocholl, M. J., Schnell, H. T., Werth, A., Herrgen, J., & Schmidt, J. E. (2016). Dialect variation influences the phonological and lexical-semantic word processing in sentences: Electrophysiological evidence from a cross-dialectal comprehension study. *Frontiers in Psychology*, *7*(739), 1–18. https://doi.org/10.3389/fpsyg.2016.00739
Leivada, E., Westergaard, M., Duñabeitia, J. A., & Rothman, J. (2021). On the phantom-like appearance of bilingualism effects on neurocognition: (How) should we proceed? *Bilingualism: Language and Cognition*, *24*(1), 197–210. https://doi.org/10.1017/S1366728920000358
Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). *emmeans: Estimated marginal means, aka least-squares means* [Computer software]. R package. https://CRAN.R-project.org/package=emmeans
Lohndal, T., & Westergaard, M. (2021). Grammatical gender: Acquisition, attrition, and change. *Journal of Germanic Linguistics*, *33*(1), 95–121. https://doi.org/10.1017/S1470542720000057

Lundquist, B., & Vangsnes, Ø. A. (2018). Language separation in bilectal speakers: Evidence from eye tracking. *Frontiers in Psychology*, *9*, 1394. https://doi.org/10.3389/fpsyg.2018.01394

Lüdecke, D. (2022). *sjPlot: Data visualization for statistics in social science* [Computer software]. R package, Version 2.8.12. https://CRAN.R-project.org/package=sjPlot

Martin, C. D., Garcia, X., Potter, D., Melinger, A., & Costa, A. (2016). Holiday or vacation? The processing of variation in vocabulary across dialects. *Language, Cognition and Neuroscience*, *31*(3), 375–390. https://doi.org/10.1080/23273798.2015.1100750

Mickan, A., & Lemhöfer, K. (2020). Tracking syntactic conflict between languages over the course of L2 acquisition: A cross-sectional event-related potential study. *Journal of Cognitive Neuroscience*, *32*(5), 822–846. https://doi.org/10.1162/jocn_a_01528

Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2015). On the left anterior negativity (LAN): The case of morphosyntactic agreement. *Cortex*, *66*, 156–159. https://doi.org/10.1016/j.cortex.2014.06.009

Navarro-Torres, C. A., Beatty-Martínez, A. L., Kroll, J. F., & Green, D. W. (2021). Research on bilingualism as discovery science. *Brain and Language*, *222*, Article 105014. https://doi.org/10.1016/j.bandl.2021.105014

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34*(6), 739–773. https://doi.org/10.1006/jmla.1995.1033

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733. https://doi.org/10.1162/089892998563121

Rodina, Y., & Westergaard, M. (2021). Grammatical gender and declension class in language change: A study of the loss of feminine gender in Norwegian. *Journal of Germanic Linguistics*, *33*(3), 235–263. https://doi.org/10.1017/S1470542719000217

Rossi, E., Kroll, J. F., & Dussias, P. E. (2014). Clitic pronouns reveal the time course of processing gender and number in a second language. *Neuropsychologia*, *62*, 11–25. https://doi.org/10.1016/j.neuropsychologia.2014.07.002

Rothman, J., & Iverson, M. (2013). Islands and objects in L2 Spanish: Do you know the learners who drop__? *Studies in Second Language Acquisition*, *35*(4), 589–618. https://doi.org/10.1017/S0272263113000387

Røyneland, U. (2009). Dialects in Norway: Catching up with the rest of Europe? *International Journal of the Sociology of Language*, *2009*(196–197), 7–30. https://doi.org/10.1515/IJSL.2009.015

Røyneland, U., & Vangsnes, Ø.A. (2020). Joina du kino imårgå? Ungdommars dialektskriving på sosiale medium. *Oslo Studies in Language*, *11*(2), 357–392.

Sandøy, H. (1988). Samsvarsbøying av adjektiv og perfektum partisipp i norske dialektar. In A. Bjørkum & A. Borg (Eds.), *Nordiske studiar: Innlegg på den tredje nordiske dialektologkonferansen* 1986 (pp. 85–117). Oslo: Universitetsforlaget.

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2015). *afex: Analysis of factorial experiments* [Computer software]. R package, Version 0.13–145. https://CRAN.R-project.org/package=afex.

Statistics Norway. (2021). *Innvandrere og norskfødte med innvandrerforeldre*. Retrieved January 14, 2022, from https://www.ssb.no/befolkning/innvandrere/statistikk/innvandrere-og-norskfodte-med-innvandrerforeldre

Tanner, D. (2019). Robust neurocognitive individual differences in grammatical agreement processing: A latent variable approach. *Cortex*, *111*, 210–237. https://doi.org/10.1016/j.cortex.2018.10.011

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, *16*(2), 367–382. https://doi.org/10.1017/S1366728912000302

Trahey, M., & White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition*, *15*(2), 181–204. https://doi.org/10.1017/S0272263100011955

Tremblay, A. (2020). *LMERConvenienceFunctions: Model selection and post-hoc analysis for (G)LMER models* [Computer software]. R package, Version 3.0. https://CRAN.R-project.org/package=LMERConvenienceFunctions

van de Meerendonk, N., Chwilla, D. J., & Kolk, H. H. (2013). States of indecision in the brain: ERP reflections of syntactic agreement violations versus visual degradation. *Neuropsychologia*, *51*(8), 1383–1396. https://doi.org/10.1016/j.neuropsychologia.2013.03.025

Zaharchuk, H. A., Shevlin, A., & van Hell, J. G. (2021). Are our brains more prescriptive than our mouths? Experience with dialectal variation in syntax differentially impacts ERPs and behavior. *Brain and Language*, *218*, Article 104949. https://doi.org/10.1016/j.bandl.2021.104949

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Accessible Summary
Appendix S1. Production Sentences.
Appendix S2. Stimuli Sentences.
Appendix S3. Statistical Models.