

MORPHOPHONOLOGICAL VARIATION IN NORWEGIAN NEGATIVE MARKER ENCLISIS

HENRIK TORGENSEN & PIOTR GARBACZ

University of Oslo

ABSTRACT

The study looks at 4 variants of negative clitics in Norwegian, how frequently they are used and which types of verbs they combine with. Using corpora of spoken Norwegian, we look at how different variants of the negative clitics vary in frequency of use and how each variant is constrained by the form of the verbs they cliticize to. In particular, we look at how the preceding vowel interacts with the negative clitics and how this interaction relates to the previous literature on Norwegian negative clitics.

[1] INTRODUCTION

The Norwegian negative markers are frequently shortened to a clitic form. In this paper, we explore what regulates this process in terms of phonology, morphology and dialectal variation. Three earlier claims about Norwegian negative marker cliticization serve as the starting point for our investigation:

- (i) Cliticization of the negative marker *ikke* is not fully productive with monosyllabic verb forms
- (ii) Cliticization of the negative marker *ikke* is not fully productive with polysyllabic verb forms
- (iii) *Ikke* may not cliticize to a preceding [ə]

Section 2 outlines the theoretical notion of clitics and introduces the literature on Norwegian negative marker cliticization. In Section 3 we use corpora of spoken Norwegian to illustrate the distribution of more variants of the negative clitic and how each variant distributes. We thus broaden the scope of the study of negative cliticization in Norwegian from the Eastern Norwegian *ikke* to include three more variants: *ikkje*, *itte* and *inte/ente*, all Norwegian variants of the English negator ‘not’. Section 4 presents the empirical baseline for our analysis. Sections

5.11–5.14 present an in-depth distribution of clitic variants and the verbs they combine with. Finally, in Section 5.2–5.4, we look at suffix elision and how negative clitics may vary more than previously thought, and present a phonological explanation for our findings. Lastly, we tie our analysis to the three original claims above.

[2] NEGATIVE CLITICIZATION IN NORWEGIAN

Both pronouns and negative markers alternate between full and clitic forms in Norwegian. Kristoffersen (2000, p. 331–339) outlines how pronouns and negative clitics exist in somewhat different syntactic environments. While pronominal clitics attach to any verbal form, preposition or subjunction, negative clitics only attach to prosodic words headed by an inflected verb. Clitics as a linguistic category challenge the basic notion of wordhood by carrying characteristics of both affixes and lexical items. Similarly to affixes, clitics are separate elements of meaning that are phonetically conjoined to another element of meaning (Booij 2012, p. 166). Yet, clitics are less restricted than affixes with respect to the word classes they attach to. To separate clitics from lexical items and affixes, Zwicky & Pullum (1985) outline several phonological and syntactic properties. Crucially, clitic forms are claimed never to appear as separate phonetic forms, and to be unable to host lexical stress. As an example, in many Norwegians dialects the 3rd singular pronoun *hun* [hʉn] – ‘she’ has a clitic variant *a* [a]. Compare (1) using the full pronominal form with the clitic form *a* in (2). The latter is ill-formed in a situation where lexical stress is used to emphasize contrast.

(1) nei HUN gjorde det
 no SHE did it

(2) *nei A gjorde det
 no SHE did it

As described by Zwicky & Pullum (1985), clitics differ from affixes in what other lexical elements they phonetically combine with. One realization of the Norwegian present tense suffix, [-ər], attaches only to verbal roots. (3a–d), taken from the Language Structure Made Accessible corpus (LIA; Øvrelid et al. 2018) and the Nordic Dialect Corpus (NDC; Johanessen et al. 2012), show how pronominal clitics may attach to inflected verbal forms (3a), prepositions (3b), nouns (3c) other pronominal clitics (3d). The former two (3a, b) use the clitic form *a*, the clitic form of *hun* as in (1) and (2), while the latter (3c, d) demonstrate two different clitic forms of *han* [han] – ‘he’; *an* and *n*.

- (3) a. det gjorde=a ikke
that did=she not
'That she didn't do.' (LIA: 1067399)
- b. han kom hit og så på=a
he came here and looked at=her
'He came here and looked at her' (LIA: 124632)
- c. så selde kommunen=an
then sold the.county=it
'Then the county sold it.' (LIA: 479197)
- d. så viste=jæ=n=det
then showed=I=him=it
'Then I showed it to him' (NDC: 3071421)

As evident from (3d), Norwegian allows for several pronominal clitics in a row. Negative markers may also take part in such clitic chains. (4) demonstrates a Norwegian phonological unit ['la:di:sɛçɛ] consisting of a verbal preterite form *la* [la:] – 'laid' followed by two pronominal clitics *di* [di] – 'they' and *se* [sɛ] – 'themselves', along with a negative clitic '*kje* [çɛ]¹, a shortened variant of *ikkje* – 'not'.

- (4) så seint ['la:di:sɛçɛ]
so late put.to.bed=they=themselves=not
'They didn't go to bed that late.' (LIA: 403197)

The last of the three clitics in (4) constitutes the topic of this article. Johannessen (1997) describes the ability of the Norwegian negator *ikke* – 'not' to cliticize to verbal roots. She notes several common contracted verb + negative marker constructions as demonstrated in (5a–g).²

- | | | | | |
|--------|---------|------|----------|--------------|
| (5) a. | har | ikke | [hakə] | 'has not' |
| b. | trenger | ikke | [trɛŋkə] | 'needs not' |
| c. | kan | ikke | [kaŋkə] | 'can not' |
| d. | skulle | ikke | [skukə] | 'should not' |
| e. | er | ikke | [ækə] | 'is not' |
| f. | står | ikke | [stokə] | 'stands not' |
| g. | får | ikke | [fokə] | 'gets not' |
- (Johannessen 1997)

[1] Also realized with a final [ə].

[2] Note the different forms '*kje* (4) and '*kke* (5a–g). As will be shown, the negative marker is realized differently according to dialect. The two forms are clitic variants of respectively *ikkje* and *ikke*, the Nynorsk and Bokmål representations of the Norwegian negative marker.

Not all verbs, however, are able to host the negative clitic. Johannessen (1997) asserts that several verbal forms are marked when followed by the negative clitic (6a–c).

- | | | | | | |
|-----|----|------|------|----------|---------------|
| (6) | a. | før | ikke | #[fœkə] | ‘feeds not’ |
| | b. | trær | ikke | #[tɾækə] | ‘threads not’ |
| | c. | ter | ikke | #[tɛkə] | ‘behaves not’ |
- (Johannessen 1997)

While the above intuitions indicate that monosyllabic verbs vary in their ability to host the negative clitic, morphologically complex verbal forms are not considered. Lindstad (1999), however, claims that when a verbal form consists of a prefix along with a verbal stem, the resulting present tense form is less able to host the negative clitic. The argument is that the morphologically complex (prefixed) forms (7b, 7d) make the cliticization of the negative marker infelicitous, while corresponding (unprefixed) monosyllabic forms (7a, 7c) are acceptable.³

- | | | | | | |
|-----|----|---------|------|-------------|-------------------------------|
| (7) | a. | tar | ikke | [takə] | ‘takes not’ |
| | b. | overtar | ikke | #[o:vɛtakə] | ‘takes over not’ |
| | c. | ser | ikke | [sɛkə] | ‘sees not’ |
| | d. | overser | ikke | #[o:vɛʃɛkə] | ‘ignores not’ [lit. over-see] |
- (Lindstad 1999)

Another and earlier account of negative marker cliticization is found in Fretheim (1988, p. 83–84), where the cliticized negative marker is claimed to be unable to combine with a final unstressed present tense marker [–ər] or the past tense marker [–ə]. This explanation does not account for the markedness of the above (7b, d), but does account for the unacceptability of the forms in (8a, b).

- | | | | | | |
|-----|----|--------|------|------------|---------------|
| (8) | a. | finner | ikke | *[finəkə] | ‘finds not’ |
| | b. | nådde | ikke | *[nåddəkə] | ‘reached not’ |
- (Fretheim 1988, p. 83)

In sum, Johannessen’s (1997) claim accounts for the acceptability of many, but not all, monosyllabic forms ending in the present tense marker [–r] such as *tar*.

Lindstad (1997) expands the claim by proposing a general constraint against negative cliticization to verbal forms whose morphological structure is

[3] Common to the examples from Johannessen (1997) and Lindstad (1999) is that the verbal forms undergo r-deletion when they host the negative clitic. [r] in the suffix [–ər] is deleted before non-coronals (5a–g) as described in Kristoffersen (2000, p. 312).

prefix+monosyllabic stem. This rules out polysyllabic verb forms such as *overtar* (7b). Many verbal forms, however, make use of the present tense marker [-ər]. These forms are always polysyllabic. Fretheim (1988) accounts for the general unacceptability of cliticizing *'kke* to these verbal forms (8a, b). The above-mentioned literature in large part, however, describes the behaviour of only one out of many variations of the Norwegian negative marker. Looking at which negators combine with which verbal forms, we have two central aims. The first goal is to investigate the claims outlined by the authors above. Explicitly, these claims are refined from the introductory claims:

- (i) Within the monosyllabic verbal forms, there is variation as to which lexical roots allow for cliticization of *ikke* (Johannessen 1997)
- (ii) Polysyllabic verbal forms derived morphologically from monosyllabic verbs do not allow for cliticization of *ikke* (Lindstad 1999)
- (iii) Morphologically underived polysyllabic verbal forms do not allow for cliticization of *ikke* when the verbal form ends in -ə(r) (Fretheim 1988)

The second goal is to document the usage of three further variants of the negative marker: *ikkje*, *itte* and *ente/inte* and to investigate how their clitic variants interact with the hypotheses above, and whether this differs from the variant *ikke*. In our investigation, we use data from two corpora of spoken Norwegian: the Language Infrastructure made Accessible (LIA) and Nordic Dialects Corpus (NDC) developed by the Text Laboratory at the University of Oslo. These corpora contain a wide variety of recordings of a wide selection of Norwegian dialects from between 1930 and 2015. The corpora are transcribed phonetically, allowing for a search of forms that deviate from orthographic standards. The two corpora differ in their years of recording with LIA containing recordings of older speakers recorded chiefly between 1957 and 1992. NDC, on the other hand, has a wider range of speaker age groups and contains recordings made between 1998 and 2015.

In the next section, we present our baseline for comparison in our study: four Norwegian negative markers, *ikke*, *ikkje*, *itte* and *ente/inte* and three present tense suffixes, -ə(r), -a(r) and -æ(r). Using corpus data, we cast new light on variation in Norwegian negative marker enclisis.

[3] VARIATION IN NEGATIVE MARKERS AND VERBAL TENSE SUFFIXES

[3.1] *Negative markers*

To investigate the broader distribution of negative cliticization across Norwegian varieties, we look at four negators. Jahr (1990, p. 21) describes five main variants of a negative marker in Norwegian. *Itte* has been the traditional form used in inland Eastern Norway, although it has nowadays (to a large extent) given way to *ikke*, in particular in urban areas (ibid.). *Ikke* is mostly associated with urban and coastal Eastern Norway, but has become common also in urban inland areas and in Finnmark. *Ikkje* is used in Western, Central, and large parts of Northern Norway as well as in the western parts of Eastern Norway. *Ente/inte* is used in coastal south-eastern areas of Norway bordering Sweden, mostly among older speakers, and has mostly replaced by *ikke* (Theil Endresen 1990, p. 98). There is considerable evidence that these four forms have shorter clitic variants. The last main variant of the Norwegian negative marker, *ittj*, used in Trøndelag, already an apocopated form descending from **int* (Dalen et al. 2008, p. 87), has no reduced form that is as readily discernible. Because of the difficulty of identifying a reduced variant of the *ittj* negative marker in the corpora, this particular variant has not been included in this search. We concede that phonetically subtler signs of cliticization in the variant *ittj* may exist, but the scope of these effects are considered too narrow for the present study.

[3.2] *Verbal present tense morphology*

Norwegian present tense forms may be monosyllabic or polysyllabic. When monosyllabic, the present tense is expressed by the suffix [-r] (9) or, in the case of auxiliary verbs, expressed by the verb root (10). As shown above, negative clitics may attach to both overtly suffixed verbal forms (5a, b, e, f, g) and unsuffixed auxiliary forms (5c).

- (9) dei står fast på det same gamle
 they stand stuck on the same old
 ‘They keep to the same old [ways].’ (LIA: 403197)

- (10) de vil fiske de er jo heilt ville
 they want to.fish they are as.you.know wholly wild
 ‘They want to fish, they are, as you know, wholly wild.’ (NDC: 2079175)

Polysyllabic present tense forms have different realizations of the tense suffix across varieties. In the present study, we limit our scope to three realizations of the present tense suffixes that attach to polysyllabic verbal forms. While the full

picture is phonetically more nuanced, we examine three common realizations of the present tense suffix. $-a(r)$ is found across urban Eastern Norway and is realized variably with or without a final $[r]$. The same holds true for the Western- and Northern Norwegian ending $-a(r)$, also found in rural Eastern Norway, and the Southeastern form $-æ(r)$. Each of the polysyllabic verbal forms have been classified according to whether it is morphologically derived from a monosyllabic verb root (7a vs. 7b) or underived (8a, b).

When the initial consonant of the cliticized negative marker is coronal, the suffix $-r$ assimilates the coronal to a retroflex, as shown in (11) and (12). Here, the $-r$ suffix has been assimilated with $[t]$ and $[n]$ respectively to form the retroflex $[ʈ]$ and the retroflex cluster $[nʈ]$.

- (11) jeg har=tte [ha:ʈɛ] gått mer enn seks og halv jeg
 I have=not gone more the six and half I
 'I haven't gone [to school] for more than six and half [years].'
 (LIA: 995801)

- (12) det var=nte[vanʈɛ] noka kjære mor
 there was=not any dear mother
 'There was no mercy [lit. dear mother].'
 (LIA: 3011145)

[4] THE INVESTIGATION

To look into which negative clitics combine with which verbal forms, we have searched the corpora mentioned above for any phonetic realization of the negative clitic that follows a present or past tense verb. The clitic forms are realized phonetically as in (13a–d). The corpora transcriptions vary somewhat in how they represent the phonetic forms. To account for this variance, different orthographic forms have been used in the search, shown in the rightmost column of (13).

- | | | |
|--------------|-------------------|--------------------------------|
| (13) a. ikke | [kə] | <ke>, <kke> |
| b. ikkje | [çə] ⁴ | <kje>, <kkje> |
| c. itte | [tə] | <te>, <tte> |
| d. ikkje | [ntə] | <nte>, <nnte>, <ntte>, <nntte> |

All four forms of the cliticized negative marker are present in the LIA and NDC corpora. (13a–d) show each cliticized negative marker following the monosyllabic present tense verb form *bli* – 'become/will be'.

[4] We thank Thorstein Fretheim (p.c.) for reminding us about further realizations of this clitic, including $[ə]$ and $[tçə]$.

(14) a. ikke	[blikə]	(LIA: 892050)
b. ikkje	[bliçə]	(LIA: 348895)
c. itte	[blitə]	(LIA: 1144621)
d. ikkje	[bliŋtə]	(LIA: 271390)

Monosyllabic verbal present tense forms are primarily important for the first of the hypotheses in Johannessen (1997). (15) displays each of the 34 monosyllabic present tense verb forms⁵ that precede any cliticized negative marker in the LIA and NDC corpora. Certain verbs whose dictionary forms are polysyllabic are occasionally realized with one syllable only. Such forms are indicated in parentheses.

- (15) blir, bor, bryr, bør, dreg, driv, er, farer (fer), faller (fell), får, gir, gjør, går, har, kan, legger (legg), må, skal, skjer, slår, ser, set, spør, står, syns, tar, tarv, tror, tykker (tykk), tør, vet, vert, vil, vær

In polysyllabic verbal forms, we investigated which of the four negative clitics combine with which of the three verbal endings. The findings are summarized in Table 1. The table indicates the total number of hits in the LIA and NDC corpora for each combination of negative clitic and verbal present case ending. For every combination of tense suffix and negative clitic, an example compound is included. The CQP string used in the search is given as (16), where the quoted text following “phon=” is subject to change to indicate each verbal⁶ and negative form. The results are limited to verbal forms between 3 and 12 letters in the orthographic convention used in the corpora. The query will naturally also include monosyllabic results.⁷ These have been excluded but will be discussed separately in Section 5.4.

- (16) [word="{3,12}" %c & pos="verb" & temp="pres" & phon=".*er"]
[phon="kke"]

[5] The verbs in (15) follow the bokmål orthographic standard of Norwegian. Because the LIA corpus is annotated following nynorsk orthography, the LIA forms have been translated directly to bokmål. In cases where a monosyllabic pronunciation contrasts a polysyllabic written form, parenthesized forms show the realization of the verb. The particular case of *tykk* is found in Vegusdal and may be analysed either as simply a monosyllabic realization of the verb or as a haplology whereby *tykkje=kje* simplifies to *tykk=kje*.

[6] It has been assumed, and to a great extent confirmed, that the orthographic form *-er* of the verbal present tense suffix corresponds to a phonetic realization *-ə(r)*.

[7] We thank Rolf Theil for having made us aware of the variants such as [çø.ø=kɛ] for *kjører ikke* ‘drives not’. We would need to investigate more such examples to determine their nature. One may assume then that the cliticization of *‘kke* is more frequent and more nuanced than the corpora indicate.

Negative marker/ verbal pres. suffix	<kke>, <ke>	<kje>, <kkje>	<te>, <tte>	<nnte>, <ntte>
-ə(r)	3 gidde=kke	84 liker=kje	7 huser=tte	10 tenker=nnte
-a(r)	25 anar=kke	324 hannlar=kje	0	2 passar=nnte
-æ(r)	12 joggæ=kke	44 kjennæ=kje	7 dansær=tte	6 husær=nnte

TABLE 1: Cliticization of different negative markers to polysyllabic verb forms according to suffix form in the LIA and NDC corpora.

Furthermore, for both corpora, we investigated the total number of clitic forms compared to the total number of full lexical forms of the negative marker. The results for each corpus are shown in Table 2, along with the ratio of clitics versus total number of negative items. This clitic ratio is a measure of how frequently each negative marker is cliticized.

Negative marker/ verbal pres. suffix	<kke>, <ke>	<kje>, <kkje>	<te>, <tte>	<nnte>, <ntte>
# full negation (LIA)	8995	11564	1860	750
# full negation (NDC)	5802	5122	2225	87
# clitic forms (LIA)	3313	12026	384	444
# clitic forms (NDC)	2811	5912	228	96
Clitic ratio (LIA)	0.27	0.51	0.17	0.44
Clitic ratio (NDC)	0.33	0.54	0.09	0.52

TABLE 2: Total number of clitic forms for each negative marker and the ratio of clitic forms to total forms of each negative marker in the LIA and NDC corpora.

For comparison, Table 3 shows the number of each present tense verb suffix in the LIA and NDC corpora. The numbers indicate whether the forms are realized with or without a final rhotic, respectively. The table demonstrates how certain realizations of the suffix appear much more frequently in the corpora than others. That is to say, if the verbal tense suffix has no impact on cliticization, we would expect a greater absolute number of clitics attached to -ər, -ə, as these endings are three times as common as the second most common verbal suffix -ar, -a. There is thus a clear imbalance between the verbal forms in terms of corpus frequency. While there is a significant imbalance between the verbal tense suffix forms, this fact has little bearing on our investigation. The infrequent verbal

endings are sufficiently attested with negative clitics and the rarity of certain other forms becomes even more conspicuous.

verbal pres. suffix	-ər	-ə	-ar	-a	-ær	-æ
# forms (LIA, NDC)	8208	22030	2790	7115	1197	359

TABLE 3: Total number of each present tense verb suffix in the LIA and NDC corpora.

[5] DISCUSSION

[5.1] *A typology of Norwegian negative contraction*

The results show that most combinations of verbal endings and negative markers exist across different varieties of Norwegian. Some combinations, however, are undocumented or rare. In this section, we review the most common combinations that exist and what constrains them morphophonologically.

Our findings allow for an estimate of how often each clitic form appears. Table 1 accounts only for the combination of polysyllabic verbal forms with cliticized negative markers. However, separating the verbs according to their number of syllables allows for an estimate of which negative markers most frequently cliticize to polysyllabic forms.

In the following, we assume that the usage of monosyllabic and polysyllabic verbs is roughly equal across the investigated variants, i.e. that the usage of a negative marker by itself does not impact the proportion of monosyllabic to polysyllabic verbs significantly. Table 4 shows the total number of clitics attaching to polysyllabic forms (Table 1) divided by the total clitics of each type, regardless of number of syllables. Table 4 is thus a measure of how frequently each negative marker combines with polysyllabic verbs. Because of the extreme relative frequency of monosyllabic verbs compared to polysyllabic verb forms, these ratios will all be low. However, there are considerable differences between the negative markers.

Negative marker	<ke>, <kke>	<kje>, <kkje>	<te>, <tte>	<nnte>, <nnte>, <ntte>
Polysyllabic clitization ratio	0.006	0.025	0.022	0.030

TABLE 4: Proportion of polysyllabic verbal hosts out of all verbal hosts for each negative clitic

The ratios in Table 4 are four to five times higher for *'kje*, *'tte* and *'nte* than for *'kke*. The former group thus cliticizes to polysyllabic verbs more frequently than *'kke*. *'kke* can likewise be assumed to more often cliticize to monosyllabic verbal forms, while the other forms have a wider range of verbs available for cliticization.

The next four sections discuss each negative marker in turn and how they distribute with regard to the morphonology of the preceding verbal forms. The findings are discussed without explicit emphasis on particular dialects, rather looking at morphophonological variation across negative markers and the tense suffix. However, this is further detailed in cases where particular geographically defined varieties behave differently under otherwise identical circumstances.

The 'kke clitic

When *'kke* cliticizes to polysyllabic forms, cliticization is more frequent when the verbal forms end in a non-schwa vowel. Cliticization of *'kke* onto verbal forms ending in *-ar* (17) or *-a* (18) is more prominent than other vowels. Twenty-five such examples were found, of which four ended in a rhotic fricative and twenty-one in a vowel. All of the rhotic fricatives are uvular [ʁ].

- (17) **likar=kke** sånn spesielt sånn gruppe eller sånn
likes=not such particular such group or such
 '[I] don't like any particular group or anything like that.'
 (LIA: 1112679)

- (18) **høra=kke** så mye på sånn
listens=not so much on such
 '[I] don't listen much to that sort of thing.'
 (LIA: 57270)

Similarly, our results indicate that the verbal ending *-æ(r)* permits cliticization of *'kke* (19). The [r] in the verbal ending consistently undergoes elision in these constructions. This is most likely because all of the investigated variants in the corpora samples combine the vowel [æ] with a coronal rhotic [r], which does not assimilate with the velar [k].

- (19) han **årkæ=kke** å dra kornlasset opp
 he **endures=not** to pull corn.load up
 'He [the tractor] does not endure pulling the load of corn up.'
 (LIA: 977448)

Schwa-final verbs allow for this sort of cliticization much more rarely. Three such data points exist in the LIA and NDC corpora. Of these three, two contain a

final uvular [ɣ] (20) and one (21) does not.

- (20) de **sier=kke** det
 they **say=not** that
 ‘they don’t say that.’ (NDC: 1380246)

- (21) jeg **gidde=kke** en gang å gjøre det
 I **endure=not** even to do it
 ‘I don’t even bother to do it.’ (NDC: 2629811)

Overall, *'kke* seems constrained by the preceding vowel in its ability to cliticize to the verb. In particular [ə] only very infrequently precedes *'kke*. This trend is possibly more pronounced outside of areas where the Norwegian rhotic is realized as an uvular [ɣ]. Comparing these results to the polysyllabic cliticization ratio of *'kke* in table 4, we might attribute the low ratio of cliticization to the rarity of schwa-final verbal forms preceding the negative marker. Similarly, however, verbal forms ending in [-æ] and [-a] might explain why the ratio is non-zero, as these verbal endings permit cliticization of *'kke* to a greater extent. With respect in particular to the low frequency of [-æ] and [-a] tense suffixes compared to [-ə] (see Table 3), the rarity of [ə] followed by *'kke* is even more notable.

Finally, it should be noted that some Northern variants do permit the combination of a coronal [r] and a velar [k]. This combination contrasts with the permeating r-deletion found in southern dialects under the same circumstances. In Northern dialects, this combination is not restricted by the preceding vowel: velar-initial *'kke* is found following both [ar] (22) and [ər] (23). What these examples may signal is that there is some dialectal variation found in the extent to which r-deletion occurs before non-coronals; southern dialects with a verb-final coronal [r] do not exhibit the same tendency. Tentatively, we note this form of cliticization is only found following monosyllabic verbal forms, and we note a more strongly articulated rhotic that resembles [r] more than [ɾ].

- (22) Vårherre **har=kke** unnt henne ein son
 Our.Lord **has=not** granted her a son
 ‘Our Lord has not granted her a son.’ (LIA: 2568579)

- (23) de **får=kke** lov å gå med hestar
 they **get=not** permission to go with horses
 ‘They are not given permission to go by horse.’ (LIA: 4232062)

The 'kje clitic

The clitic *'kje* is more widespread than *'kke* in cliticizing to polysyllabic verbs (see Table 1, 4). Together with *'nte*, *'kje* has the highest ratio of cliticized forms to full lexical forms. 52 percent of all instances of this negative marker are cliticized forms, while 48 percent are full forms. This aligns with a previous study of negative clitic frequency of *'kje* in the Stryn dialect, Munch (2013, p. 168), where 45% of negators surface as clitic forms. Unlike *'kke*, which forces elision of coronal rhotics in the verbal suffix, *'kje* freely combines with both uvular (24), and coronal rhotics (25) and vowel-final verbal forms (26). In comparison to the clitic *'kke* (apart from Northern dialects), *'kje* thus displays no universal r-deletion (Kristoffersen 2000, p. 312), even though it is non-coronal.

- (24) det [komɑr]=[çə] høyare huset enn det første
 it comes=not higher the.house than the first
 'This house will not be higher than the first.'
 (NDC: 288526)

- (25) det [sønɑr]=[çə] folk som kjem ifrå andre kanter
 that understands=not people who come from other places
 'People who come from other places don't understand that.'
 (NDC: 3038357)

- (26) nei eg [tɛŋkə]=[çə] så mye på det
 no I think=not so much on it
 'No, I don't think about it so much.'
 (LIA: 1757996)

The 'tte clitic

Varieties of Norwegian that use the negative marker *itte* may cliticize it as *'tte*. Starting with a coronal [t], this form undergoes assimilation with the present tense suffix of the verbal form. In (26), a rhotic consonant is assimilated to a coronal [t]. While the overall frequency of cliticization of this form is lower than the other negators in Table 4, it has a substantially higher frequency of cliticizing to polysyllabic verbal forms like (26) than the other plosive-initial clitic, *'kke*. Only enclisis to the verbal suffixes *-ar* (26) and *-ær* (27) were found in our corpora.

- (27) dem [dɔnsæ]=[tə] nå som dem gjorde før i tida
 they dance=not now like they did before in time
 'They don't dance like they used to now.'
 (LIA: 996191)

While no polysyllabic verbal forms ending in *-ar* are found in our material, many

instances of the monosyllabic *har* (28) or *skal* (29) precede *'tte*. Thus, there are seemingly no constraints on the vowel that precedes *'tte*.

- (28) jeg [hɑ]=[tə] tatt gulrot
I have=not taken carrot
'I haven't stolen [any] carrots.' (LIA: 992624)

- (29) [ska]=[tə] jeg si noe om
should=not I say anything about
'I shouldn't say anything about that.' (LIA: roemskog_03gm)

What our results gently hint at is that the usage of the *'tte* clitic has declined over time compared to the other clitic variants. While all other cliticization ratios are stable across the older corpus (LIA) and the newer corpus (NDC), the ratio of *'tte* to *itte* is halved from 0.17 to 0.09 (Table 3). While this decline is outside the scope of this article, it falls in line with trends of regionalization of Norwegian dialects, and the spread of Urban East Norwegian (Mæhlum & Røyneland 2012, p. 52–53).

The 'nte clitic

The clitic form *'nte* is frequently used among those speakers who use the negative marker *inte/ente*. Forty-five percent of all instances of this negator are cliticized forms. Out of all clitic forms, *'nte* is the most frequent in cliticizing to polysyllabic verbal forms. Similarly to *'tte*, *'nte* assimilates preceding rhotics into a retroflex cluster [ŋ], following *-ar* (31), *-ar* (32) and *-ær* (33). Furthermore, all results give the assimilated consonant cluster, with no elision of the rhotic [ɾ].

- (31) det [hœra]=[ŋtə] til det stellet
that belongs=not to that place
'That does not belong that that place.' (LIA: 3006913)

- (32) men du [kʁɛmə]=[ŋtə] hardt
but you squeeze=not hard
'But you don't squeeze hard.' (LIA: 251432)

- (33) passert sytti at dem [orkæ]=[ŋtə] like mye lenger
passed seventy that they bother=not as much anymore
'[When they] pass seventy years old, they do not bother as much any-
more.' (NDC: 692971)

[5.2] *Elision of the present tense suffix – are both 'kje and ikkje possible clitic forms?*

In consonant-final verbal forms like *tull-er* 'joke-s' the present tense suffix vowel [-ə~a~æ] is sometimes entirely assimilated to the negative marker. Thus, the vowel following the verbal root may be realized as [ɪ] (34). When, additionally, the /r/ undergoes deletion, the cliticized complex is left without an overt present tense suffix, reducing the polysyllabic verbal form to a single syllable identical to the lexical verb root. In (34), the verb is represented only by the verbal root *tull*, with no tense suffix. We should interpret such compounds as separate from verbs that lack overt present tense suffixes, because the reduced verbal form does not exist when the negative marker is lexically stressed, such as (35). The variant that is ruled out is the one that combines lexical stress and the reduced verbal stem (36), because the reduced verbal root is invoked by cliticization itself, but lexically stressed negators do not cliticize.

- (34) du **tull=ikkje**[tuliçə] med farfar liksom
 you **mess=not** with grandpa kinda
 'You kinda don't mess with grandpa.' (NDC: 2577767)

- (35) du **tullar IKKJE** med farfar liksom
 you **mess not** with grandpa kinda
 'You kinda DON'T mess with grandpa.' (modified from (34))

- (36) *du **tull IKKJE** med farfar liksom
 you **mess not** with grandpa kinda
 'Intended: You kinda DON'T mess with grandpa.' (modified from (34))

This raises the question of how tense is communicated in constructions when the lexical root is the only realized verbal element. A tentative explanation might be found by exploring whether past tense suffixes have the same property of being phonologically elided as the present tense marker. Regardless, our data hints toward an analysis where both the full form and the reduced form of the negative marker may cliticize to the verb. The reduced form attaches to the verbal root+suffix compound as in (31, 32, 33) and the full form attaches to a reduced verbal root as in (34).

The above division of the negative marker clitics holds true for negative markers that do not assimilate with the present tense suffix. Neither 'kje nor 'kje phonologically assimilate to the final rhotic in the present tense suffix. 'tte and 'nte, however, mostly undergo retroflexion of the initial coronal as in (37) or (38).

- (37) en [gɪdə]=[ɲtə] å reise bort og skyte
 one bothers=not to go there and shoot
 'One does not bother to go there and shoot.' (LIA: 155052)

- (38) jeg [føɾjə]=[tə] med
 I follow=not with
 'I do not pay attention (lit. *follow with*).' (LIA: 984258)

A similar argument for two negative clitics is found in data from Fretheim (1988, p. 83). Similarly to (34) and (35), negative clitics freely attach to past tense forms as in (39), but the past tense suffix [-də] is realized only by [d]. When the negative clitic is lexically stressed, however, the full past tense suffix [-də] is obligatory, leaving (40) well-formed, but (41) marked. The minimal difference between (40) and (41) is the presence of [ə] within past tense suffix [-də].

- (39) jeg [nəd]=[ɪkə] fram
 I reached=not forth
 'I couldn't get there.' (own)

- (40) jeg [nədə] IKKE fram
 I reached NOT forth
 'I couldn't get there [emphasis on not].' (own)

- (41) *jeg [nəd] IKKE fram
 I reached NOT forth
 'Intended: I couldn't get there [emphasis on not].' (own)

As for clitic forms of the negative marker that begin with a coronal, such as '*n*te, we predict that if elision of the tense prefix happens in varieties using '*n*te or '*t*te negative clitics, the retroflexion of the initial coronal consonant in the negative marker will not occur. This is because it is the suffix that mediates the retroflexion. The hypothetical result is along the lines of (42). The realization of such examples is particularly important to determine whether the suffix is actually elided; an elided suffix can be expected not to trigger retroflexion.

- (42) en [gɪd]=[mtə] å reise bort å skyte
 one bothers=not To Go there and shoot
 'One does not bother to go there and shoot.' (modified from (37))

As we primarily examined verbs that have an overt tense suffix, examples of the (42) are sparse in our data. We cannot at this stage make any concrete claims about the distribution or behavior of suffix elision compared to *normal*

cliticization (i.e. a negative marker cliticized to an overt tense prefix). Yet, the identical behavior of past tense verbs (39, 40, 41) and suffix-elided present tense verbs (34, 35, 36) under lexical stress presents an interesting case for the theoretical aspect of Norwegian clitics. Further, the two datasets raise the question of why suffix elision (33) happens instead of simple enclisis (30–32). A more comprehensive investigation into the geographical distribution of suffix elision as a result of negative cliticization is invited. The exact geographical distribution of the phenomenon is difficult to ascertain from our dataset.

In the next section, we look into explanations for why certain negative clitics occur more frequently with polysyllabic verbs.

[5.3] *Phonological explanations*

Certain combinations of verbal endings and negative markers are highly limited. In this section, we examine which combinations allow for cliticization, and provide explanations for why the distribution is as different as it is.

The most prominent result of our query is the disproportional rarity of the clitic *'kke* following polysyllabic verbal forms compared to the other negative clitics. Within the *'kke* clitics, however, there is some variation between which present tense suffixes in polysyllabic verbs allow for cliticization. While both *-a(r)* and *-æ(r)* may precede the *'kke* clitic, *-ə(r)* is much more limited in this capacity, in particular when speakers realize the rhotic as a coronal rather than as a uvular.

Our findings indicate a tendency in the initial consonant of the negative clitic towards non-velars ([ç], [n], [t]) when the present tense suffix is realized as *-ə(r)*. This finding can be accounted for in terms of articulation – cliticization is easier if the verb-final [r] is able to assimilate to the initial consonant of the negative clitic. Similarly, this accounts for the possibility of speakers who use uvular rhotics [ʁ~ʀ] to cliticize *'kke* and *'kje*, to the verb. For speakers using the coronal rhotic [r], both suffix vowels [ə] and [a] are limited to one hit each. The latter group, those using the present tense suffix *-a(r)* might be explained through morphological ambiguity. [r] does not assimilate with [k], and is omitted in such combinations. This would leave the past and present tenses of verbs such as *huske* ‘remember’ ambiguous when the negative marker is cliticized: both the past and present tense would be realized as [hʊskakə]. While such forms are ambiguous, there is no general ban on [a] preceding the negative clitic *'kke*: all of the monosyllabic forms *har'kke* [hakə] ‘has not’, *var'kke* [vəkə] ‘was not’ frequently surface.

For speakers using the suffix vowel [ə], however, no such ambiguity would arise from a hypothetical cliticized form *[hʊskəkə]. The same goes for speakers

using the suffix vowel [æ] to form [hʊskækə], a form present in several speakers. These results point towards some property of [ə] that is not present in [a] or [æ], disallowing it preceding the negative clitic. Our findings fall in line with the rule proposed by Fretheim (1988, p. 83), which disallows [ə] before [k] in Eastern Norwegian. The exact nature of this property lies beyond the scope of this paper, but one might tentatively assume that it is related to intensity or duration. A future phonetic study might invite an answer to this particular question.

[5.4] *Revisiting Johannessen (1997) and Lindstad (1999)*

Returning to the previous claims about Norwegian cliticization, we first look at polysyllabic derived verbs. Lindstad's (1999) claim is that more morphologically complex verbal forms are infelicitous in cliticized contexts. In our data, two morphologically complex verbal forms, *forstår* 'understands' and *behøver* 'needs', are particularly frequent in combination with negative clitics. The disyllabic verb *forstå* is on its own is found 514 times in the data from LIA and NDC. Twenty-nine of these 514 instances are followed by the negative clitic (42, 43). Likewise, the trisyllabic *behøver* appears 140 times in the corpora, 16 of whose are encliticized forms (44, 45, 46). Six of the 16 hits are reduced to disyllabic forms (45), while a further 6 are reduced to monosyllabic forms (46).

The prosodic behaviour of the two verbs *forstå* and *behøve* must be considered, however. Both verbs are loanwords where stress is assigned respectively finally and prefinally (Bokmålsordboka). Because of the final stress, *forstår* thus behaves prosodically as a monosyllabic verb, and we would in fact expect it to allow for cliticization. *Behøver* carries prefinal stress and an overt present tense ending [-ər], and as such behaves as an ordinary polysyllabic form. We may thus attribute the felicitous cliticization of the negative marker onto these verbs as a consequence of their stress assignment.

- (43) eg **forstår=kke** det at dei kunne bygge så stort
 I **understand=not** it that they could build such large
 'I don't understand how they could build such a large [house].'
 (LIA: 2649691)
- (44) **forstår=nte** at det gikk an
understand=not that It went on
 '[I] don't understand how it was possible [lit. went on].'
 (LIA: 2861936)
- (45) det **behøve=kje** vare
 that **needs=not** last
 'That doesn't have to last.'
 (NDC: 481378)

- (46) **behø=kke** å ta hensyn til andre folk
 needs=not to take considerations to other people
 ‘[one] does not need to take others into consideration.’ (LIA: 4074736)
- (47) da **bøv=kje** eg visa dykk nedi kjellaren
 then needs=not I show you down.in the.basement
 ‘But then I don’t have to show you the basement.’ (LIA: 957513)

The exact verbal forms presented in Lindstad (1999), *overtar* and *overser*, however, are much rarer at 7 and 1 hits respectively across both corpora. As opposed to *forstår* and *behøver*, these verbs carry initial stress. None are followed by the negative marker in any corpus. A tentative search of 10 more morphologically derived verbs (48) yielded only a single occurrence of any verb being followed by a full negative marker. No negative clitics were found. In total, the 10 verbs yielded a total of only 18 hits, illustrating their rarity across the board (compare 514 instances of *forstår*).

- (48) avgår, avser, avtar, gjentar, inngår, innser, inntar, overgår, oppnår, tilber

Because of their general rarity, then, we find it difficult to empirically judge Lindstad’s claim. We might however speculate about what sort of data might surface if there is, contrary to Lindstad, no morphological constraint on cliticization of the negative marker. In the principal investigation in this study, a strong bias against the combination of [-ə(r)] and [k] was found. If this phonological constraint, but no morphological constraint, holds for (48), we would expect that the negative clitics ‘*kje*, ‘*tte* and ‘*nte* freely cliticize to every verb. The clitic ‘*kke*, however, would be limited to those verbs not ending in [-ə(r)]. This is the purely phonological hypothesis.

Looking again at the class of verbs proposed as uncliticizable by Lindstad because of their morphology, we could equally well postulate that they are unable to host clitics because of the prosody inherent to this class of verbs. Lindstad’s example verbs all have in common that stress is realized somewhere before the final syllable. Typologically, this stress pattern does not correlate with a ban on cliticization. As shown, several varieties of verbal endings and negative markers are possible across Norway with disregard to the distance between primary stress and the cliticized negative marker. The only broad generalization found in our data is the constraint against [ə(r)k]. If prosody is not, then, the factor that disallows stress, the most salient interpretation of Lindstad’s argument is indeed that the speaker is aware of the morphology that underlies

every verb. Lindstad's argument thus reduces to a mental awareness of prefixes and roots for individual verbs. As shown above, our empirical counterclaim is that only a phonological constraint *[ə(r)k] is needed, and that the available data is too sparse to posit a morphologically grounded constraint. We predict that forms such as *overser=kke* or *bidrar=kke* would be found given a larger database, and that these forms are not ruled out because of morphology. Further, there might be some degree of variation in cliticization frequency across areas using the *ikke* forms, where some varieties allow for more clitic-bearing hosts. Looking beyond morphology, we might postulate independent reasons that forbid cliticization: word length or word frequency.

Importantly, hypothetical forms such as *innser=kke* combined with a constraint against cliticization to normal (i.e. not morphologically complex) polysyllabic verbs such as **fniser=kke*. Such forms would raise further questions about the actual phonetic realization of the final vowel. It seems uncontroversial to assume that the final vowel of *innser* resembles more closely [e(:)] than it resembles the final [ə] vowel of *fniser*.

Looking now to the monosyllabic verb class, we recorded 34 different present tense verb forms cliticizing with a negative marker. These are repeated in (49). The parenthesized forms indicate a monosyllabic realization of a polysyllabic dictionary form.

- (49) blir, bor, bryr, bør, dreg, driv, er, farer (fer), faller (fell), får, gir, gjør, går, har, kan, legger (legg), må, skal, skjer, slår, ser, set, spør, står, syns, tar, tarv, tror, tykker (tykk), tør, vet, vert, vil, vær

On its own, this list would seem to support Johannessen's claim that some monosyllabic verbs do not cliticize to the negative marker. The particular verbs *før* 'feeds' and *ter* 'behaves' mentioned in Johannessen (1997) are, however, rare in the LIA and NDC corpora. Both verbs yield one hit each across both corpora. Neither of these are negated in the context. Looking at polysyllabic forms, speakers using the clitics *'kje* and *'nte* exhibit no obvious constraints on the polysyllabic verbs that are cliticized onto. The extreme productivity of cliticization in polysyllables might simply indicate that polysyllabic forms are more abundant in the speakers' lexicons than monosyllabic forms. We see no obvious reason why the same productivity would not be present in monosyllabic verbs, however. If a certain monosyllabic form is used frequently, we might predict that it is also cliticized more productively.

Our counterclaim regarding monosyllabic cliticized forms is thus that frequency determines possibility of cliticization. The particular forms ruled out as ungrammatical by Johannessen are not ruled out for arbitrary reasons. Our

counterclaim would explain the ruling out of forms such as *før=kke* [føkə] (Johannessen 1997, p. 4) by referring to its rarity. Many modern speakers who only rarely, if ever, use this verb, might not easily be able to cliticize the negative marker onto it. Yet, we posit that speakers who use the verb frequently would also be able to cliticize the negative marker onto it. We hypothesize that the same constraint rules out the combination *ter=kke* for a majority of modern speakers.

[6] CONCLUSION

We have investigated how Norwegian negative markers cliticize to verbs across different realizations of negative markers and verbal endings. We examined three previous claims about Norwegian negative marker enclisis: two morphological hypotheses and one phonological. Using data from two corpora of spoken Norwegian, we have shown how certain combinations of verbal suffixes and negative markers are disproportionally frequent or infrequent. Thus we have aimed to broaden the typological scope and theoretical understanding of Norwegian negative marker enclisis, branching out from the focus on the *ikke* variant of the negative marker. We account for our distribution using arguments motivated by articulatory closeness, while also suggesting that there might be some phonological property of the Norwegian [ə] that makes it less common than [a] or [æ] preceding negative clitics.

We further examined the case of suffix elision, whereby the tense suffix is completely elided and a negative clitic attaches to the verb root. Lexical stress was used as a diagnostic to show some of the peculiarities of this process, indicating the negative clitic may have more elaborate forms than previously considered.

TAKK

Takk til de to anonyme fagfellene samt til Øystein A. Vangsnes for givende kommentarer og synspunkter på et tidligere utkast av denne artikkelen. Kommentarene har bidratt til vesentlig forbedring! Alle resterende feil og mangler er helt og fullt forfatterens ansvar.

Med denne artikkelen vil vi gjerne hedre Janne, som indirekte har gjort dette arbeidet mulig gjennom å ha ledet Tekstlaboratoriet og egnet store deler av sitt forskningsliv til å bygge forskningsinfrastruktur som vi nå kan benytte. Som leder for Tekstlaboratoriet og som sjef var Janne en inspirasjon på flere måter: Det var et storartet sosialt miljø som fikk de nyansatte til å virkelig like å ta del i prosjektet. Janne har gjennom å inkludere dusinvis av studenter i forskningsarbeidet ved UiO fostret flere forskere som nå er virksomme i hele

Norge og utenlands. Hun har også på enestående vis vist hvordan professorer kan (og bør) være en stor positiv kraft for studentene sine. og hun står som eksempel for andre og til inspirasjon for de neste. Takk, Janne, for alt du har vært og alt du har gitt oss!

REFERENCES

- ‘Behøve’, ‘Forstå’. In: *Bokmålsordboka*. Språkrådet og Universitetet i Bergen. <http://ordbok.uib.no> (retrieved 06.10.20).
- Booij, Geert. 2012. *The grammar of words: An introduction to linguistic morphology*. Oxford University Press.
- Dalen, Arnold. 2008. Trøndermåla ca. 1950–1850. In *Trøndersk språkhistorie*, edited by Arnold Dalen, Jan Ragnar Hagland, Stian Hårstad, Håkan Rydving & Ola Stemshaug, 19–82. Trondheim: Tapir akademisk.
- Endresen, Rolf Theil. 1990. Vikværsk. In *Den store dialektboka*, edited by Ernst Håkon Jahr, 89–99. Novus forlag.
- Fretheim, Thorstein. 1988. Den kontraherte negativpartikkelen: ‘ke eller ‘kke. *Norsk Lingvistisk Tidsskrift*, 6, 73–88.
- Jahr, Ernst Håkon. 1990. *Den store dialektboka*. Novus forlag.
- Johannessen, Janne B. 1989. Klitika – en avgrensning. *Norsk Lingvistisk Tidsskrift*, (2): 117–147.
- Johannessen, Janne B. 1997. *The status of the negation ikke in Norwegian*. Ms. University of Oslo.
- Johannessen, Janne B., Joel Priestley, Kristin Hagen, Anders Nøklestad & André Lynum. 2012. *The Nordic dialect corpus*. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation. European Language Resources Association*, edited by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Ugur Dogan, Bente Maegaard, Joseph Mariani, Jan Odijk, 3388–3391.
- Kristoffersen, Gjert. 2000. *The phonology of Norwegian*. Oxford: Oxford University Press.
- Lie, Svein. 1990. Indre Østlandet. In Jahr (1990), 101–119.
- Lindstad, Arne M. 1999. *Issues in the syntax of negation and polarity in Norwegian. A minimalist analysis*. Cand. philol. thesis. University of Oslo.

- Munch, Christine B. Ø. 2013. *North Germanic negation. A microcomparative perspective*. Cand. philol. thesis. University of Tromsø.
- Mæhlum, Brit & Unn Røyneland. 2012. *Det norske dialektlandskapet: innføring i studiet av dialekter*. Oslo: Cappelen Damm Akademisk.
- Øvrelid, Lilja, Andre Kåsen, Kristin Hagen, Anders Nøklestad, Per E. Solberg & Janne B. Johannessen. 2018. *The LIA treebank of spoken Norwegian dialects*. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*, edited by Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis & Takenobu Tokunaga, 4482–4488. European Language Resources Association.

CONTACTS

Henrik Torgersen
Universitetet i Oslo
hatorger@uio.no

Piotr Garbacz
Universitetet i Oslo
piotr.garbacz@iln.uio.no