

NB Uttale Work on Dialects

The aim of this project has been to add dialectal transcriptions to a pronunciation lexicon in order for it to better represent dialects beyond Eastern Norwegian. To do so, we have divided the Norwegian dialects into five broad dialect areas and identified features that can be considered representative for each area. In addition to Eastern Norway, the areas are Southwestern Norway, Western Norway, Central Norway and Northern Norway. To arrive at the representative features we want to include in the lexicon, we have had to weigh certain aspects against each other.

There is great dialectal variation within each of the five dialect areas and we cannot cover all the variation. It is beneficial to cover dialectal features and phenomena that are representative of a large group of speakers, which often means referring to dialects spoken in and around larger cities. At the same time, it is advantageous if the pronunciation variants differ from Eastern Norwegian and the other dialect areas, so that the lexicon captures the greatest possible variety. This, however, often suggests dialects spoken by fewer speakers and the variants could be disappearing, which makes them less relevant for the lexicon. We have weighed these considerations against each other and chosen dialectal features that are both typical of the area and have a certain prevalence.

We have also had to consider whether it is possible to add the dialectal variants using regular expressions. The rulesets use regular expressions to find and change phoneme sequences in a group of words that share a phonemic pattern. It is also possible to target words based on grammatical features or orthography. Dialect phenomena that are unpredictable and inconsistent within word groups are therefore difficult to include in the lexicon. (An example of this is lenition of unvoiced plosives in Southern and Southwestern Norwegian, which is no longer a productive process. It is not possible to distinguish the lenited plosives based on phonemic environment, orthography, or grammatical features).

We have relied on available literature on Norwegian dialects to determine what dialectal features to include in the lexicon, mainly *Dei norske dialektane* by Martin Sjøkkeland and *Den store dialektboka* by Ernst Håkon Jahr, as well as other relevant articles and theses. [Store norske leksikon's](#) entries on Norwegian dialects have served as a useful summary of the most important features that distinguish Norwegian dialects. Furthermore, we have used the speech corpora [Nordic Dialect Corpus](#) and [NB Tale](#) as sources for information on pronunciation in spontaneous speech.

The literature emphasises morphology, specifically the inflections of nouns and verbs, as important dividing lines between dialects. We have therefore had the same focus on grammatical categories within this project. Moreover, grammatical inflections are well-suited for the type of rule-based approach this project has, seeing as we can formulate rules that match inflectional suffixes common to a group of words that share grammatical features.

We have chosen to refer to specific dialects within each dialect area as a kind of reference dialect. The dialects have been selected on the grounds that their inflectional paradigms differ to some extent from the other dialect areas, that the dialectal features are representative of a larger group of speakers, and that it is possible to find literature about the dialects as well as examples of pronunciation. The five dialects we have chosen are all described in the literature and represented in the speech corpora.

For Northern and Central Norway, we have chosen the dialects spoken in the cities Bodø and Trondheim, respectively. An important phenomenon in both Northern and Central Norway is apocope. Apocope is present in both these dialects but behaves slightly differently. In the Salten area (Bodø), apocope differentiates between a-verbs and e-verbs, while in Trondheim, apocope is a result of vowel balance in Old Norse. Both dialects also have vowel lowering, although this feature is more typical of Northern Norway.

For Southwestern Norway, we have chosen to work with the Jæren dialect. The dialect is representative of a large group of speakers as the Jæren area includes the cities Sandnes and Stavanger. The Jæren dialect belongs to southwestern *a-mål* and has some *nynorsk* typical features.

For Western Norway, we have chosen to refer to the dialect in Indre Sogn. We have chosen not to use Bergen as a reference dialect, even though it is a dialect spoken by many speakers in and around Bergen city. This is because the dialect is atypical for Western Norway as a whole and has a number of similarities with Eastern Norwegian that would not result in further variants in the lexicon.

The dialect in Indre Sogn belongs to the northwestern *a-mål* and thus has some features in common with Southwestern Norwegian. However, the dialect also has features that are typical for Sogn og Fjordane and Sunnmøre. The *a-mål* is not as consistent across inflectional categories, and the inflections differ somewhat from Southwestern Norwegian. The dialect also has pronunciation of plosives that are mute in other dialects.

Errorfixes and dialectal variants

Before we started the work of adding dialectal variants to the pronunciation lexicon, we implemented some errorfixing rules to correct transcription errors in the existing Eastern Norwegian lexicon. The names of these rulesets begin with *errorfix*. An important part of this stage was to correct the wrong use of retroflex phonemes before adding transcription variants where the retroflexes are broken up into the rhotic and following sound. The lack of retroflex pronunciations, where the rhotic and following sound are pronounced separately, is a distinctive feature of both Southwestern and Western Norwegian.

The rulesets that add dialectal variants to the lexicon have names beginning with *dialect*. The dialect updates can be divided into five umbrella categories:

- Breaking up retroflex phonemes
- Noun inflections
- Verb inflections
- Adjective inflections

- Other vowel and consonant features

The lexicon has two transcription categories for each of the five dialect areas: a spoken category and a written category. Most of the dialect rules are only implemented for the spoken category, as this is meant to represent pronunciations in spontaneous speech. The written category reflects pronunciations closer to the written standard of *bokmål*, and therefore remains unchanged for most of the updates. The exception is non-retroflex pronunciations in Southwestern and Western Norwegian since the process of retroflexion does not occur in these dialects at all.

In the lexicon, the dialect areas are represented by letters (e, n, sw, t, w) followed by the transcription category:

- Eastern Norway: e_spoken, e_written
- Northern Norway: n_spoken, n_written
- Southwestern Norway: sw_spoken_sw_written
- Central Norway: t_spoken, t_written
- Western Norway: w_spoken, w_written

Appendix A lists examples of transcriptions and how these transcriptions are changed from Eastern Norwegian to the other dialects. For further details about the work on dialect updates, including descriptions of dialectal features and pronunciation variants, see the Norwegian documentation *NB Uttale Dialektarbeid*.

Appendix A

The appendix contains examples of words and transcriptions for each dialect area and shows what updates have been implemented. The transcriptions following an arrow → show the resulting dialectal transcription for a group of words with the same inflectional suffix or phonemic pattern in that dialect. Examples that do not have a second transcription, illustrate that the transcription is the same for that dialect as already existed in the lexicon and no update has taken place.

The examples are ordered according to the five categories of dialect updates mentioned above. The examples contain the Norwegian word, original transcription, updated transcription, English translation, and grammatical features.

Breaking up retroflex phonemes

Examples of breaking up the retroflex phoneme /RN/:

fiskeren, /F IH2 S K AX0 RN/ → /F IH2 S K AX0 R AX0 N/ (*the fisherman*), noun
trenernes, /T R EE2 N AX0 R RN AX0 S/ → /T R EE2 N AX0 R N AX0 S/ (*the coaches*'), noun

Noun inflections

Central Norwegian

Masculine nouns

hest, /H EH1 S T/ (*horse*), sg. (singular) ind. (indefinite)
fisker, /F IH2 S K AX0 R/ (*fisherman*), sg. ind.
hesten, /H EH1 S T NX0/ (*the horse*), sg. def. (definite)
fiskeren, /F IH2 S K AX0 RN/ (*the fisherman*), sg. def.
hester, /H EH2 S T AX0 R/ → /H EH2 S T AH0/ (*horses*), pl. (plural) ind.
fiskere, /F IH2 S K AX0 R AX0/ → /F IH2 S K AX0 R AH0/ (*fishermen*), pl. ind.
hestene, /H EH2 S T NX0 AX0/ → /H EH2 S T AH0 N/ (*the horses*), pl. def.
fiskerne, /F IH2 S K AX0 RN AX0/ → /F IH2 S K AX0 R AH0 N/ (*the fishermen*), pl. def.

Feminine nouns

jente, /J EH2 N T AX0/ (*girl*), sg. ind.
jenta, /J EH2 N T AH0/ (*the girl*), sg. def.
jenter, /J EH2 N T AX0 R/ → /J EH2 N T AH0/ (*girls*), pl. ind.
jentene, /J EH2 N T NX0 AX0/ → /J EH2 N T AH0 N/ (*the girls*), pl. def.

Neuter nouns

hus, /H UU1 S/ (*house*), sg. ind.



National Library of Norway

huset, /H UU1 S AX0/ (*the house*), sg. def.

hus, /H UU1 S/ (*houses*), pl. ind.

epler, /EH2 P L AX0 R/ → /EH2 P L AH0/ (*apples*), pl. ind.

husa, /H UU1 S AH0/ (*the houses*), pl. def.

eplene, /EH2 P L AX0 N AX0/ → /EH2 P L AH0 N/ (*the apples*), pl. def.

Northern Norwegian

Masculine nouns

hest, /H EH1 S T/ (*horse*), sg. ind.

fisker, /F IH2 S K AX0 R/ (*fisherman*), sg. ind.

hesten, /H EH1 S T NX0/ (*the horse*), sg. def.

fiskeren, /F IH2 S K AX0 RN/ (*the fisherman*), sg. def.

hester, /H EH2 S T AX0 R/ → /H EH2 S T AH0/ (*horses*), pl. ind.

fiskere, /F IH2 S K AX0 R AX0/ → /F IH2 S K AX0 R AH0/ (*fishermen*), pl. ind.

hestene, /H EH2 S T NX0 AX0/ → /H EH2 S T AH0 N/ (*the horses*), pl. def.

fiskerne, /F IH2 S K AX0 RN AX0/ → /F IH2 S K AX0 R AH0 N/ (*the fishermen*), pl. def.

Feminine nouns

jente, /J EH2 N T AX0/ (*girl*), sg. ind.

jenta, /J EH2 N T AH0/ (*the girl*), sg. def.

elver, /EH2 L V AX0 R/ → /EH2 L V AH0/ (*rivers*), pl. ind.

jenter, /J EH2 N T AX0 R/ → /J EH2 N T AX0/ (*girls*), pl. ind.

elvene, /EH2 L V AX0 N AX0/ → /EH2 L V AH0 N/ (*the rivers*), pl. def.

damene, /D AA2 M AX0 N AX0/ → /D AA2 M AX0 N/ (*the ladies*), pl. def.

jentene, /J EH2 N T NX0 AX0/ → /J EH2 N T NX0/ (*the girls*), pl. def.

Neuter nouns

hus, /H UU1 S/ (*house*), sg. ind.

huset, /H UU1 S AX0/ (*the house*), sg. def.

hus, /H UU1 S/ (*houses*), pl. ind.

epler, /EH2 P L AX0 R/ → /EH2 P L AH0/ (*apples*), pl. ind.

husa, /H UU1 S AH0/ (*the houses*), pl. def.

eplene, /EH2 P L AX0 N AX0/ → /EH2 P L AH0 N/ (*the apples*), pl. def.

Southwestern Norwegian

Masculine nouns

hest, /H EH1 S T/ (*horse*), sg. ind.

fisker, /F IH2 S K AX0 R/ → /F IH2 S K AH0 R/ (*fisherman*), sg. ind.

hesten, /H EH1 S T NX0/ (*the horse*), sg. def.

fiskeren, /F IH2 S K AX0 RN/ → /F IH2 S K AH0 R AX0 N/ (*the fisherman*), sg. def.

hester, /H EH2 S T AX0 R/ → /H EH2 S T AH0/ (*horses*), pl. ind.

fiskere, /F IH2 S K AX0 R AX0/ → /F IH2 S K AH0 R AH0/ (*fishermen*), pl. ind.

hestene, /H EH2 S T NX0 AX0/ → /H EH2 S T AH0 N AX0/ (*the horses*), pl. def.

fiskerne, /F IH2 S K AX0 RN AX0/ → /F IH2 S K AH0 R AH0 N AX0/ (*the fishermen*), pl. def.

Feminine nouns

jente, /J EH2 N T AX0/ → /J EH2 N T AH0/ (*girl*), sg. ind.

jenta, /J EH2 N T AH0/ → /J EH2 N T OAH0/ (*the girl*), sg. def.

jenter, /J EH2 N T AX0 R/ → /J EH2 N T AX0/ (*girls*), pl. ind.

jentene, /J EH2 N T NX0 AX0/ (*the girls*), pl. def.

Neuter nouns

hus, /H UU1 S/ (*house*), sg. ind.

huset, /H UU1 S AX0/ (*the house*), sg. def.

hus, /H UU1 S/ (*houses*), pl. ind.

husa, /H UU1 S AH0/ → /H UU1 S OAH0/ (*the houses*), pl. def.

Western Norwegian

Masculine nouns

hest, /H EH1 S T/ (*horse*), sg. ind.

hesten, /H EH1 S T NX0/ (*the horse*), sg. def.

hester, /H EH2 S T AX0 R/ → /H EH2 S T AH0/ (*horses*), pl. ind.

hestene, /H EH2 S T NX0 AX0/ → /H EH2 S T AH0 N AX0/ (*the horses*), pl. def.

Feminine nouns

jente, /J EH2 N T AX0/ → /J EH2 N T AH0/ (*girl*), sg. ind.

jenta, /J EH2 N T AH0/ (*the girl*), sg. def.

bygda, /B YH1 G D AH0/ → /B YH1 G D IH0/ (*the village*), sg. def.

jenter, /J EH2 N T AX0 R/ → /J EH2 N T AX0/ (*girls*), pl. ind.

jentene, /J EH2 N T NX0 AX0/ (*the girls*), pl. def.

Neuter nouns

hus, /H UU1 S/ (*house*), sg. ind.

huset, /H UU1 S AX0/ (*the house*), sg. def.

hus, /H UU1 S/ (*houses*), pl. ind.

husa, /H UU1 S AH0/ → /H UU1 S IH0/ (*the houses*), pl. def.

Verb inflections

Central Norwegian

Regular verbs

håpe, /H OAH2 P AX0/ → /H OAH2 P/ (*hope*), inf. (infinitive)

angre, /AH2 NG R AX0/ (*regret*), inf.

fikle, /F IH2 K L AX0/ (*fidget*), inf.



National Library of Norway

banker, /B AH2 NG K AX0 R/ → /B AH2 NG K AX0/ (*knocks*), prs. (present)
dømmer, /D OEH2 M AX0 R/ → /D OEH2 M AX0/ (*judges*), prs.

kjøpte, /KJ OE2 P T AX0/ → /KJ OE2 P T/ (*bought*), pst. (past)

Irregular verbs

leser, /L EE1 S AX0 R/ → /L AE1 S/ (*reads*), prs.
sover, /S OA1 V AX0 R/ → /S OE1 V/ (*sleeps*), prs.
bærer, /B AE1 R AX0 R/ → /B AE1 R/ (*carries*), prs.
holder, /H OAH1 L AX0 R/ → /H EH1 L/ (*holds*), prs.

funnet, /F UH2 N AX0 T/ → /F UH2 N AX0/ (*found*), pst. part. (participle)
ligget, /L IH2 G AX0 T/ → /L OH2 G AX0/ (*lain*), pst. part.

Northern Norwegian

Regular verbs

håpe, /H OAH2 P AX0/ → /H OAH2 P/ (*hope*), inf.
angre, /AH2 NG R AX0/ → /AH2 NG AX0 R/ (*regret*), inf.
fikle, /F IH2 K L AX0/ → /F IH2 K LX0/ (*fidget*), inf.

banker, /B AH2 NG K AX0 R/ → /B AH2 NG K AH0/ (*knocks*), prs.
dømmer, /D OEH2 M AX0 R/ → /D OEH2 M/ (*judges*), prs.

kjøpte, /KJ OE2 P T AX0/ → /KJ OE2 P T/ (*bought*), pst.

Irregular verbs

leser, /L EE1 S AX0 R/ → /L AE1 S/ (*reads*), prs.
sover, /S OA1 V AX0 R/ → /S OE1 V/ (*sleeps*), prs.
bærer, /B AE1 R AX0 R/ → /B AE1 R/ (*carries*), prs.

funnet, /F UH2 N AX0 T/ → /F UH2 N AX0/ (*found*), pst. part.
ligget, /L IH2 G AX0 T/ → /L OH2 G AX0/ (*lain*), pst. part.

Southwestern Norwegian

Regular verbs

håpe, /H OAH2 P AX0/ → /H OAH2 P AH0/ (*hope*), inf.

banker, /B AH2 NG K AX0 R/ → /B AH2 NG K AH0/ (*knocks*), prs.
dømmer, /D OEH2 M AX0 R/ → /D OEH2 M AX0/ (*judges*), prs.

Irregular verbs

gir, /J II1 R/ → /J EE1 R/ (*gives*), prs.
gjør, /J OE1 R/ → /J EE1 R AX0/ (*does*), prs.

holder, /H OAH1 L AX0 R/ → /H EH1 L AX0/ (*holds*), prs.

sover, /S OA1 V AX0 R/ → /S OE1 V AX0/ (*sleeps*), prs.

har, /H AA1 R/ → /H EE1/ (*has*), prs.

Western Norwegian

Regular verbs

håpe, /H OAH2 P AX0/ → /H OAH2 P AH0/ (*hope*), inf.

dekkes, /D EH2 K AX0 S/ → /D EH2 K AH0 S T/ (*cover*), passive inf.

banker, /B AH2 NG K AX0 R/ → /B AH2 NG K AH0/ (*knocks*), prs.

dømmer, /D OEH2 M AX0 R/ → /D OEH2 M AX0/ (*judges*), prs.

Irregular verbs

gir, /J II1 R/ → /J EE1 R/ (*gives*), prs.

gjør, /J OE1 R/ → /J EE1 R AX0/ (*does*), prs.

holder, /H OAH1 L AX0 R/ → /H EH1 L AX0/ (*holds*), prs.

sover, /S OA1 V AX0 R/ → /S OE1 V AX0/ (*sleeps*), prs.

Adjective inflections

Northern Norwegian

kortere, /K OAH2 RT AX0 R AX0/ → /K OAH2 RT AH0 R/ (*shorter*), comparative

spennende, /S P EH2 N NX0 AX0/ → /S P EH2 N AH0 N AX0 S/ (*exciting*), prs. part.

Southwestern Norwegian

spennende, /S P EH2 N NX0 AX0/ → /S P EH2 N AH0 N AX0 S/ (*exciting*), prs. part.

Western Norwegian

spennende, /S P EH2 N NX0 AX0/ → /S P EH2 N AH0 N D AX0/ (*exciting*), prs. part.

Other vowel and consonant features

The updates in this group are not tied to grammatical categories or inflectional paradigms but are rather changes to vowels and consonants in certain phonemic or orthographical environments, or they are lexically conditioned.

Central Norwegian

veldig, /V EH2 L D IH0/ → /V AEH2 L D IH0/ (*very*), adverb

hjem, /J EH1 M/ → /J AEH1 M/ (*home*), adverb

spille, /S P IH2 L AX0/ → /S P AEH2 L AX0/ (*play*), verb

fryktelig, /F R YH2 K T AX0 L IH0/ → /F R OEH2 K T AX0 L IH0/ (*horrendous*), adjective

fisk, /F IH1 S K/ → /F EH1 S K/ (*fish*), noun

Northern Norwegian

veldig, /V EH2 L D IH0/ → /V AEH2 L D IH0/ (*very*), adverb

hjem, /J EH1 M/ → /J AEH1 M/ (*home*), adverb

spille, /S P IH2 L AX0/ → /S P AEH2 L AX0/ (*play*), verb

fryktelig, /F R YH2 K T AX0 L IH0/ → /F R OEH2 K T AX0 L IH0/ (*horrendous*), adjective

fisk, /F IH1 S K/ → /F EH1 S K/ (*fish*), noun

mitt, /M IH1 T/ → /M EH1 T/ (*mine*), possessive pronoun

forelder, /F OAH0 R EH1 L D AX0 R/ → /F OEH0 R EH1 L D AX0 R/ (*parent*), noun

utenfor, /UH2 T NX0 F OAH3 R/ → /UH2 T NX0 F OEH3 R/ (*outside*), preposition

for, /F OAH1 R/ → /F OEH1 R/ (*for*), preposition

fare, /F AA2 R AX0/ → /F AE2 R AX0/ (*go*), verb

fått, /F OAH1 T/ → /F OAH1 D/ (*gotten*), verb pst. part.

Western Norwegian

sang, /S AH1 NG/ → /S AH1 NG G/ (*song*), noun

ringe, /R IH2 NG AX0/ → /R IH2 NG G AX0/ (*ring*), verb

strand, /S T R AH1 N/ → /S T R AH1 N D/ (*beach*), noun

under, /UH1 N AX0 R/ → /UH1 N D AX0 R/ (*under*), preposition

kveld, /K V EH1 L/ → /K V EH1 L D/ (*evening*), noun

ilden, /IH1 L NX0/ → /IH1 L D NX0/ (*the fire*), noun