Reduction of alternative pronunciations in the Norwegian computational lexicon NorKompLeks

Torbjørn Nordgård and Arne Kjell Foldvik

Department of Linguistics
Norwegian University of Science and Technology
Torbjorn.Nordgard@hf.ntnu.no
arne.foldvik@hf.ntnu.no

Abstract

This paper describes a method for selecting one single pronunciation from a set of alternatives. Two types of reductions are described – base form reductions and inflected form reductions. The strategies are different because the inflected forms have to take care of two properties of Norwegian: The language allows for many spelling alternatives, both in base forms and inflected forms, and many morphological inflection types license pronunciation alternatives. A comprehensive pronunciation lexicon with more than 65000 base forms is used, and it is demonstrated that it is possible for users of the lexicon to derive their own stylistic preferences.

1. Introduction

The computational lexicon NorKompLeks contains, among other things, information on pronunciation for all officially recognized word forms in Norwegian ([1], [2]). The lexicon has approximately 65000 bokmål base forms and 90000 nynorsk ones. (Bokmål and nynorsk are the two officially recognized orthographies of Norwegian.) All base forms are manually controlled by senior phoneticians. Inflected paradigms are generated from the base forms, and all words in the paradigms are phonologically described. Since the inflected forms are machine generated, they are not manually controlled to the same extent as the base forms, even though of course tests have been carried out.

2. Pronunciation Variation

The pronunciation database of NorKompLeks mirrors the variation in Southeast Norwegian pronunciation (roughly the greater Oslo area), typically listing more than one pronunciation alternative. For automatic speech recognition it is important that the lexicon gives a wide coverage of pronunciation variants, especially for a language like Norwegian where dialect pronunciations are encountered in all social contexts. No widely accepted pronunciation norm exists. In NorKompLeks the alternatives of relaxed, colloquial Southeast Norwegian pronunciation are listed according to assumed frequency, with the most common pronunciation listed as no. 1, etc. The pronunciation alternatives relate mainly to two aspects: Placing of stress and word tone and phoneme realizations.

The listing of alternative pronunciation in decreasing order of frequency is based on intimate knowledge of Southeast Norwegian pronunciation among the project workers. The descriptions of the pronunciation are in x-sampa [3], which is a keyboard compatible coding of all IPA symbols and diacritics [4].

2.1. Stress and word tone.

There is a phonologically interesting pattern of stress / word tone in words of foreign origin. Many words of this type, e.g. bensin (fuel) and stasjon (station) have two different pronunciations in Southeast Norwegian: one with word tone 2 and stress on the first syllable, the other with word tone 1 and the main stress on the last syllable. Less common among words of foreign origin are the pronunciation alternatives found with some words like massiv (massive): Either tone 1 and stress on the first syllable or main stress on the last syllable (and tone 1).

2.2. Variations in phoneme realizations.

The other phonologically distinguishing source is variations in phoneme realizations. An example is the use of the retroflex flap versus the apical tap. This concerns words with the spelling 'rd', like bord (table), which is pronounced either with a retroflex flap finally, bu:rl, or with an alveolar tap, bu:r in Southeast Norwegian. The retroflex flap also alternates with the dental/alveolar lateral in many words written with 't', thus the T in blå (blue) is pronounced either as a retroflex flap, brLO:, or as a lateral, bLO:.

3. Reducing the variations in base forms

For speech synthesis purposes the ordering of pronunciation according to assumed frequency in relaxed, colloquial speech is not altogether unproblematic. First of all the ordering according to alleged frequency may not be flawless, and secondly the use or non-use of e.g. retroflex flap instead of tap or lateral, indicates stylistic variation. Variation of this type should not be mixed accidentally in speech, be it human or machine generated, since it would signal bad command of spoken Norwegian.

During the autumn of 2000 the authors initiated a project with the aim of reducing the variation of pronunciation in a way that would facilitate specifications of phonological parameters for different synthetic speech purposes. The aim was to provide a lexicon with one and only one pronunciation per word form, and the lexicon user should be able to make his own decisions regarding pronunciation style.

In carrying out this project, we were faced with two problems: First, reducing the number of alternative pronunciations in the basic word forms, then performing a similar reduction in the inflected word forms. The two tasks
appeared to be quite different, for reasons we will describe below.

3.1. Reductions in base forms

For the basic word forms a set of general parameters were defined, facilitating removal of a pronunciation with property X if an alternative pronunciation, property Y, is provided for the same word form, and all other phonological properties in the competing descriptions are equal. An example is that the pronunciation 'gi:' is given preference over 'Si:'. This will affect all word forms where this variation is provided, e.g. in a word like energi (energy), which has the following pronunciation description in NorKompLeks:

1. energi eSn{r$'gi:, eSn{$'Si:

After removal of Si: we have

2. energi eSn{r$'gi:

The reduction only affects the two phonemes 'g' and 'S', but it is important that the right context 'i:' is provided here. Absence of relevant contexts may lead to unexpected reductions, and the work indeed demonstrated that relevant contexts should be provided whenever possible.

Another example concerns T, which can be given preference over 'rL'. This is a much more frequent phenomenon than the 'gi:' / 'Si:' alternation. An example is kule (bullet)

3. kule "k}:$l@, "k}:$rL@

After removal we arrive at

4. kule "k}:$l@

when the pronunciation with 'rL:' is removed.

Variations in stress placement were mentioned earlier. This reduction is special in the sense that it appears to be quite systematically connected to tone variation, not only in the empirical realm but also in the XSAMPA notation. Stress and tone are marked simultaneously by the two symbols " and "", where "" means stress and tone 1 whereas """" means stress and tone 2. So the question is whether one should reduce tone variation or stress placement. As mentioned in 2.1 there are cases where stress can be placed at different locations with the same tone. For this reason we decided to use stress placement as the relevant parameter.

3.2. Reduction ordering

When reductions of the type described above are applied the question of reduction ordering becomes relevant because a word may have more than one candidate for reduction. Situations like this were anticipated when the pronunciations were made, and, consequently, each potential variation in a base form is described separately, yielding two descriptions for one variation, four variations for two variations, and so on. This pattern is adjusted by the phoneticians when appropriate, which means that there are cases with e.g. three descriptions for a given word form.

Since the pronunciations are given in this fashion in the lexicon, the reductions can be applied in arbitrary order. We have not observed any empirical effects so far which can be attributed to reduction orderings. It is, however, natural to apply the most pervasive reductions first in order to make the reduction process efficient.

The syllable internal reductions are as follows in the currently used reductions offered in the NorKompLeks lexicon:

- l rL
- r rL
- @n n_.=
- @l l_.=
- @n m_.=
- s S
dr tr
- Or O:
nAd nA:
2.r 2:
r@n @rn
- vr r
- Or O
t@r t@
- AN An
- l@n @rn
- r-d rd
- n2.r j2.r
- bluts blu:s		# Tone 2 over tone 1
- t{k tAk
- vArmt vArnrt
- " "
v@
- kOn kun
- O u
- "Sl sl
ste:d ste:
grAs grA:s
- puNt puNkt
- A:rt Art
- SO SA
- mAN men
- pAN pen
- sAN sen
d2:d d2:
da:g dA:
gi: Si:
b}n bOn
- f{r-d f{r

Crucially, by switching the ordering of the pairs, other reductions will apply. Lexicon users can make their own preferences by editing the relevant text file.

3.3. Pronunciation styles

It is of course an interesting empirical question whether the results of the reductions above are acceptable in all circumstances. It must be emphasized that “acceptable” in this context means “coherent pronunciation style”. The pronunciations in isolation are all acceptable, given the fact that the base forms from which they are chosen are manually
controlled. Preliminary tests indicate that this method seems to give good results.

3.4. Syllable borders

Syllables are marked in the lexicon, and phoneme descriptions with syllables occasionally violate the requirement that reductions can apply only when the surrounding contexts are identical in the competing pronunciation descriptions. Consider the lexeme *handeļ* (trade):

5. handeļ "hAn$\ddot{d}@l","hA\$n@l"

Here the relevant phoneme sequences are ‘n’ and ‘nd’. It would not be appropriate to suppress ‘n’-variants in favor of ‘nd’, both because there is a syllable border intervening between ‘n’ and ‘d’, i.e. ‘n$\ddot{d}$’, but also because the left contexts are different: ‘‘hA’ and ‘‘hA$\$’’. Consequently, the entire phonological sequence including syllable markers must be provided: ‘n$\ddot{d}$’ outscores ‘Sn’.

3.5. Multi-phonemic patterns

The lexeme *orden* (order) has the alternations r-d@n and rdn_ which are reduction targets. The latter has a syllabic ‘n’ as kernel:

6. orden "Or-$\ddot{d}@n","Osrdn_=

Here we observe that syllable boundaries intervene again, and they must be included in the reduction specifications. In practice we have to allow for arbitrary long reduction patterns. The general observation is that pronunciation alternatives across syllable boundaries are not very uniform in character, but seem to tie up with types of phoneme strings which seem to depend heavily on lexemes. This is not a problem, but a powerful and necessary property of the method because we will often want to refer to entire lexemes as reduction targets, simply because Norwegian, like Swedish, Dutch, German etc., is a language which allows for productive morphological compounding.

In the current version of the lexicon 116 reductions with syllables are defined. The list could have been made longer by including reduction patterns with just a few occurrences, but the main reason is that pronunciation alternatives is more conveniently handled individually.

3.6. The algorithm for base form reduction

The algorithm which is used to perform the reductions of base forms is quite simple, cf. figure 1. The interpretations of the conditions and functions assumed in the algorithm should be easy to grasp and implement. The repeat loop must be adjusted for cases where no reduction pattern is available. In our implementation we have relied on the orderings made in the original lexicon (recall that the most frequent variant is listed first). A reasonable alternative would be to list all such cases and perform a control for each of them.

OR EACH base form *base* in the lexicon DO
IF *base* has multiple pronunciation(s) *prons* THEN
REPEAT UNTIL *cardinality*(prons) = 1
FOR EACH *p* in *prons* DO
IF a reduction pattern *rp-select* matches *p* THEN
IF *prons* contains *p’* where *rp-reduce* matches AND
context(*p*) = context(*p’*) THEN
delete *p’*
FI
FI
EF
FI
F

Figure 1 Base form reduction algorithm

4. Reduction of inflected forms

When it comes to inflected forms the reduction in pronunciation alternatives is more complicated. To some extent this is due to the many alternative inflected orthographic forms allowed in Norwegian (on the morphological level), but the main reason is that the pronunciation of inflectional endings can vary. An example of a general variation is the use or absence of syllabic nasal ‘n_’- in words like ‘bilen’ (=the car) and ‘bilene’ (=the cars).

Obviously, base form reduction must apply prior to morphological and phonological inflection.

4.1. Inflection in NorKompLeks

Before going into the details of how pronunciation reduction is accomplished in inflected forms, some remarks on the inflection system in NorKompLeks are necessary. Each morphological base form has a set of morphological inflection codes. These codes generate the correct morphological inflection paradigms. It turns out that a morphological inflection can be connected to different phonological realizations. As an illustration we use the variation between syllabic n (‘n_’-’) and ‘’@n’, which is quite pervasive for masculine nouns. It is also used in present and adjectival participles for many verbs (but then with a final ‘’@’’).

The noun *bil* (car) has the morphological code m1 (the most common masculine inflection). This code simply adds the suffixes –en (the car), –er (cars) and –ene (the cars) to the base form. There are however as many as eight distinct phonological realizations of this code. Only a few are applicable for a given word. Our example word *bil* selects codes which associates “bi:$l@n and “bi:$Sn_ = with the singular definite form *bilen* and ““bi:$l@$Sn@ and ““bi:$Sn_=$@ with the plural definite form *bilene*. The two singular forms occur because the base form has a final ‘l’, and the plural forms with tone 2 are generated because the base form is monosyllabic. Polysyllabic base forms must not be assigned tone shift in the plural forms, as in e.g. *lastebil* (truck), which has tone 2 with stress on the first syllable in the singular and plural forms.

4.2. Reducing inflected pronunciation variation

In reducing pronunciation variation in the inflected paradigms the applicable phonological codes must be identified for the
chosen pronunciation of the base form. From these codes one candidate is selected, and this selection can be governed by the users of the lexicon. In our example with bil (car) two phonological inflection codes are relevant; one of them providing endings with syllabic ‘n’ (code m1_fon1n) and the other with ‘@n’ (dubbed m1_fon1). One of the two has to be selected.

The syntax of the reduction descriptions can be illustrated by a couple of examples:

- m1_fon1, m1_fon1n = m1_fon1
  # “bi:$l@n,” bi:l$n_ = (the car)
- m1dt_fonrn, m1dt_fon = m1dt_fon
  # “s2s$t@$r@n, “s2s$t{rn (the sister, masculine form)

We identified 101 reductions of this type for the entire lexicon.

4.3. The algorithm for inflected form reduction

The algorithm for inflected form reduction can be stated informally as follows: For each phonological base form (base form reduction has applied) and for each of its morphological inflection codes, select the relevant phonological inflection codes. Given the set of phonological codes, consult the reduction descriptions for the set and select the decided code. Then apply the code to the phonemic base form.

4.4. Problems with inflected form reduction

During our work with the inflected form reductions we experienced that the many alternative inflected orthographic forms allowed in Norwegian, make it quite difficult to select one alternative from a set of possibilities in a coherent way. For a discussion of this phenomenon, see [5]. This has to do with the interaction between morphological and phonological variation, which in some cases leads to many descriptions of lexemes used in isolation or in compounds. Selecting from the inflected pronunciation of lexemes of foreign origin, turned out to be particularly difficult because such words occasionally have very unusual phonological properties.

A more serious issue is the fact that if the selection preferences for base forms are changed, this choice might affects the selection regime connected with the inflection codes. This is not by itself a problem for our method, but rather a linguistic phenomenon which will arise for any systematic description of phonological inflection with the aim of selecting from sets of alternatives.

5. Discussion

The issue of reducing pronunciation variation is far from trivial. If it is to be done manually, it is indeed very labor intensive. Our approach is a mixture in the sense that it requires careful manual identification of phonological patterns and, afterwards, automatic selection from each of these patterns, given selection decisions made by the users of the lexicon. We believe that our method makes it possible to adjust the lexicon to distinct pronunciation styles in a principled and empirically sound way. The number of reduction descriptions is approximately 250, which is quite reasonable given the complexity of pronunciation reduction.

6. Conclusions

The results obtained in this work is relevant for other pronunciation dictionaries for at least two important reasons. First, we have demonstrated that a phonologically sound and efficient method exists for pronunciation reduction. And second, since the inflection machinery of NorKompLeks can be compiled into e.g. finite state transducers or other inflection tools, we are convinced that the insights from this work can be used in connection with other phonological lexicons as well.

7. Acknowledgments

We would like to thank Telenor Research and Development for financial support to the work reported here.

8. References