Understanding standard language: A psycholinguistic look at Danish regional and casual speech variation

Nicolai Pharao

University of Copenhagen, Denmark

Written language is standardised to a high degree in many countries across the world. Aspects of spelling, grammar and lexis are officially normed at the national level as in France or Denmark, or through widespread accepted use of a particular dictionary as the prescriptive norm, as in Austria or the United Kingdom. Few countries (if any) have a similarly codified norm for spoken language, particularly for native speakers, but norms for proper and correct ways of speaking still exist in most (if not all) speech communities, and are enforced among members of the speech community whether explicitly or implicitly. This is certainly the case in Denmark; and the norms for standard speech in Denmark are the focus of this chapter. A series of comprehensive studies of language attitudes among Danes conducted by Tore Kristiansen and colleagues provide a detailed look at the ideologies surrounding regional speech variation. These studies are briefly reviewed here, as they provide an important background for the method I propose and exemplify in this chapter: the use of psycholinguistic tools to study standard language. Building on a small number of studies of word recognition (Ernestus , Bayeen and Schreuder 2002; Floccia et al. 2006; Sumner et al. 2014), I examine both regional variation in Danish as well as reduction that is the result of spontaneous speech processes, presumably common to all varieties of Danish. Reduction of words, such as the weakening or lack of segments as compared to a distinct pronunciation, e.g. vocalisation or even complete loss of post-vocalic /l/ in words like *whole, feel, mile*, has received some attention in experimental phonetics and psycholinguistics. However, reduction is rarely the explicit object of study in language attitude studies, and has not been purposely included in the Kristiansen studies, which focus on regional variation. But reduction in general is a part of the debate about proper speech in Denmark, or rather about 'bad' speech, as commentaries and debates almost invariably focus on speakers whose Danish is "bad" or "ugly". A common criticism in these proscriptive debates is that the speech of certain groups is "sloppy and indistinct" and "difficult to understand" because sounds or syllables "get swallowed". (Regional variation is also sometimes criticised for being difficult to understand, of course, but not necessarily because they are perceived as sloppy.) Therefore, both reduction and regional variation may be considered as relevant dimensions in what constitutes the

norm for standard spoken Danish; and the proposal in this chapter is that, by studying both at the same time, we may gain better insight into implicitly held norms than by continuing the tradition of studying either spontaneous speech process or regionally based variation.

LANGUAGE ATTITUDES IN DENMARK

In a series of studies, Tore Kristiansen and colleagues have studied the language attitudes of Danish speakers in five different locations in Denmark (summarised, in greater detail than is possible here, in Grondelaers and Kristiansen 2013 and in even greater detail in Kristiansen 2009). A notable feature of this series of studies is that both methods involving direct measures of language attitudes as well as methods using indirect measures are used. Because the two methods reveal a fundamental difference in the patterns with striking consistency across locations, this has led Kristiansen to posit the existence of both an overt and covert norm for speech variation in Denmark.

According to the results of the direct method of a Label Ranking Task, in which 15 – 16 year old participants were presented with a list of names of Danish 'ways of speaking' (including *rigsdansk* as the name for the 'Standard' together with dialect names covering the whole country) and were asked to rank them in order of the best kind of Danish to the worst kind of Danish, participants across the five locations consistently rated their own local accent or that associated with the nearest city as the best. Overtly then, young Danes prefer their own local variety.

However, in the Speaker Evaluation Experiments, a different pattern emerges. The same participants who had performed the Label Ranking Task in the five locations were presented with approximately 30 seconds of speech produced by (1) two male and two female speakers of conservative Copenhagen-based Danish, (2) two male and two female speakers of modern Copenhagen-based speech, and (3) two male and two female speakers of their own local accent. Where the conservative and modern accents differ in segmental properties (mainly vowel qualities) the local accents differ from the Copenhagen-based accents mainly in terms of prosodic features. So the set of Copenhagen-based stimuli was identical in all five locations, but the stimuli with local speech varied between locations. The design was a verbal guise experiment, with all stimuli containing descriptions of what it means to be a good teacher. Participants were asked to rate each speaker on a set of eight scales representing personality traits commonly used in language attitudes research (for details see Grondelaers and Kristiansen 2013; Kristiansen 2009). Again, with striking consistency, Kristiansen and colleagues found that the speakers representing the local accent are rated the lowest on all of the eight scales in all five locations. This

of course means that Copenhagen-based speech is consistently rated the highest. Further, the conservative variety is rated higher on semantic differential scales which may be classified as traits of superiority (such as being goal-directed or conscientious) and the modern variety is rated highest on scales which can be classified as traits relating to dynamism (such as fascinating and cool). Covertly, then, young Danes have a higher regard for speakers from Copenhagen than for speakers from their own area.

Thus, if 'standard Danish' is to be defined in terms of what seems – on empirical grounds – to count as 'best language', as argued by Kristiansen, the answer to the question what is standard Danish is not a straightforward one. In overt discourse about language, the results suggest a preference for local varieties of speaking, suggesting the existence of multiple local or regional standard Danishes. But in terms of the evaluation of speakers, Copenhagen-based language always wins out. This is interesting in the present context because, as Kristiansen has argued and as supported by a large-scale study of language change in real time reported in Maegaard et al. (2013), the covert norm is predictive of language change across several regional speech communities, whereas the overt norm is not.

These studies all look directly at the ideological dimensions of speech variation to assess what for language users themselves counts as the best language, and thus what may be deemed standard Danish in the minds of Danes. The link to language change is particularly important, because it suggests a link between attitudes and language use.

In this chapter, I present the results of a series of psycholinguistic experiments that explore the implications of the existence of such evaluations of different ways of speaking for how listeners understand spoken words. Understanding is meant here in the sense of the (so-called) low level process of recognising and hence decoding spoken words (rather than the more complex process of comprehension of the content). A small body of recent studies in psycholinguistics have examined such processing of spoken words and found a benefit for distinct and standard word forms: they are recognised faster and may even facilitate recognition of the same word pronounced with reduced segments or regional segmental variants (Sumner and Samuel 2005, 2009). That is, language users do not simply prefer some variants of word forms over others in the sense that they hold speakers who use them in higher esteem, but they also find them easier to recognise. This is in spite of the fact that in language use, distinct word forms are not the most common forms encountered in everyday interaction, nor do speakers of regional accents encounter speakers of the standard variety most often.

The series of psycholinguistic studies reported here are based on the design in Sumner and Samuel (2009) in their study of U.S. English. Here, I use Danish to explore the role of 1) reduced segments, 2) regional segmental variation and 3)

regional prosodic variation in the recognition of spoken words by the same groups of listeners. By including easiness of recognition as a criterion for defining the standard variety, and examining all of these three aspects of speech variation and their effect on the same listener population, we get a more nuanced picture of what counts as standard for Danes, and what the implications may be in everyday interaction.

LEXICAL DECISION TASKS INVOLVING REDUCTION, REGIONAL VARIATION AND PROSODY

To study the role of variation in immediate decoding of the speech signal in terms of both distinctness and regional variation, lexical decision tasks involving two segmental variables in Danish were conducted in sequence. The reduction of intervocalic /g/ and the realisation of the suffix /əð/ were used as variables. Intervocalic /g/ may be realised as either a stop or an approximant, e.g. the word /lɔgə/ *lukke* ('close' vb.) has the distinct realisation [lɔg̊ə] and the reduced realisation [lɔɰə]. The reduced realisation of intervocalic /g/ is more frequent than the distinct variant in running speech (Pharao 2011). I will therefore refer to intervocalic /g/ as the reduction variable, because the variation concerns levels of distinctness in spontaneous speech, which are relevant in all varieties of Danish. Note that it is not because [ɰ] is an approximant, that it is classified as reduced, but because it is an approximant *that is a variant of a stop*. The classification thus rests on segmental typologies in phonetics and phonology where segments are described as weaker the more sonorous they are, but crucially in this context when they are more sonorous than the corresponding segment in the word form they can reasonably be argued to be derived from within the same variety. In other words, this classification rests on the assumption that the segment [ɰ] is part of the inventory of allophones in Danish but not a phoneme in the language and that the form with [g̊] is the canonical (or underlying) form. One purpose of the experiments reported here is to study how this conceptualisation of the representation of word forms relates to the way in which listeners process these word forms. Given that the reduced variant is the more frequent one in disyllabic words of the type used in this experiment, a usage-based approach would predict that [lɔɰə] would be recognised either quicker than or as quickly as [lɔg̊ə].

The realisation of the suffix /əð/ is conditioned by region: it is realised as [əð̩] in Copenhagen Danish, but variably as either [əð̩] or [əd̥] in Aarhus Danish, a regional variety of Danish spoken in western Denmark. For example, /hʌbəð/ *hoppet* 'jumped' is realised as [hʌb̥əð̩] in Copenhagen Danish, but as [hʌb̥əd̥] in Aarhus Danish (Jensen and Maegaard 2012). I will refer to this as the regional variable,

because the variants are associated with the two regions represented here. No claim is made about reduction with respect to this variable, i.e. the Copenhagen variant [əˌð] is not hypothesized to be weaker than the Aarhus variant [əˌd], because the approximant is not a variant of an underlying /d/ (outside of hyper-distinct pronunciations as discussed above). Indeed, from a usage-based perspective, the variant [əˌd] should take longer to process for listeners from Copenhagen than the variant [əˌð] because it does not occur in their variety.

Experimental design

Participants

Two groups of listeners were recruited for the experiment: 32 from Copenhagen (mean age 22.6; 11 male, 21 female) and 32 from Aarhus (mean age 24.9; 11 male and 21 female). All listeners in the Copenhagen group were born and raised in Greater Copenhagen and lived in Copenhagen at the time the study was conducted. For the Aarhus group, all were born and raised in Eastern Jutland, where Aarhus is the biggest town, and 20 of the participants lived in Aarhus at the time of the study. The remaining 12 had moved to Copenhagen to go to university, but had only lived in Copenhagen for 18 months or less. Including two groups of listeners allowed us to test whether regular exposure to the regional variant facilitates recognition of spoken words containing this variant.

Stimuli

The stimuli for the first phase of the experiment were produced by two female speakers, one from Copenhagen and one from Aarhus. Since the purpose was to test the effect of the segmental variation alone, the tonal stress group pattern had to be kept constant in the regional variable, as the tonal stress group patterns differ between Copenhagen and Aarhus Danish, and the regional segmental variant only occurs in Aarhus Danish, at least outside of hyper-distinct "spelling pronunciation" which occurs rarely and is outside the style of speech that the experiment is attempting to study. Note that for some speakers in Copenhagen, the suffix can be realised with a stop if the stem contains and intervocalic /ð/. For example, the word /bæːðə/ (bathe) in the preterite takes the suffix /əð/. This may be realised as either [bæːˌðˌð] (with schwa-assimilation to the approximant) or [bæːˌðəˌd] where the approximant of the suffix is dissimilated from the approximant of the stem. Because the realisation with a stop is still an option (although probably mostly prevalent among older speakers in Copenhagen), words containing /ð/ in the stem were not included as stimuli.

For the reduction variable, listeners heard tokens produced in their own accent, i.e. Copenhagen listeners heard them with a Copenhagen-based tonal contour and Aarhus listeners heard them with an Aarhus-based tonal contour. Again, this was done to focus on the segmental variation for each group of participants. As a consequence of this, the first phase of the experiment was divided into two blocks, one for each segmental variable. Accent differed between the two blocks for Copenhagen participants, but remained the same for Aarhus participants.

Procedure

For the first phase of the experiment, participants simply conducted a lexical decision task involving either reduction or regional variation. The stimuli were blocked by variable, that is, participants first heard a block of 22 items involving reduction (interspersed with 30 fillers and 30 non-words, randomised for each listener) and then a block of 22 items involving regional variation (again interspersed with 30 fillers and 30 non-words, randomised for each listener). The order of the two blocks was reversed for half of the participants in each group to control for order effects. All critical items and fillers were matched for word frequency in running speech in Danish (i.e. they all fell within the mid-frequency range of words in running speech as assessed using the LANCHART corpus of sociolinguistic interviews, cf. Pharao 2011). Stimuli were presented auditorily over headphones via a laptop running the software Zep (Veenker 2018). In the beginning of the experiment, on-screen written instructions asked participants to listen to a series of words and nonsense words, and for each one to press the key on the keyboard labelled green, if they thought it was a word in Danish, and to press the key labelled red if they thought it was a nonsense word. Stimuli were presented only once and with an inter-stimulus interval of 3000 milliseconds. The experiment began with a set of 12 unrelated items to familiarise participants with the task.

Results of the lexical decision tasks

Response times were measured from stimulus offset. Raw response times to real words were analysed after removing all data points more than two standard deviations from the mean in order to reduce skew, thereby allowing for the use of mixed effects linear regression models in the analysis of the data.[1]

[1] It should be noted that for all of the statistical tests reported in this chapter, the regression models were also fit to log transformed response times, which further reduces skewness in the data. No significant differences were found between these model fits and those fit to raw response times, so the results for the raw response times are reported here, as they are more directly interpretable.

The reduction variable: Intervocalic /g/

A full model including variant and listener origin as predictors as well as presentation order, word frequency and hand preference (whether the participant was left-handed or right-handed) as fixed effects and individual listener as random effect was fit to the response time data for the critical items of the reduction variable. Only items that received a correct response were included. This full model was then stepped down, and only the significant factors are reported in the final model summary.

Table 1: Model summary for the reduction variable.

Factor	*Estimate*	*Std. Error*	*t value*	*Pr(>|t|)*
(Intercept)	455	30.5	14.8	0.0019 **
Variant: reduced	98	10.1	9.6	<2e-16 ***

Table 1 shows that, for the reduction variable, only the factor 'variant' emerged as significant ($p < 0.001$), with the reduced variant having an estimate of 98 (std. error 10.1), i.e. listeners on average took 98 milliseconds longer to recognise reduced word forms as compared to distinct word forms. Note that this holds for participants from both locations, as listener origin did not emerge as a significant factor. The error rates for words with the reduction variable also show a small but significant effect of variant ($p < 0.001$), with words with reduced /g/ having an error rate of 15% as compared to 2% for words with distinct /g/. In summary, listeners took longer to recognise words with reduced /g/ and were also more likely to classify them as non-words than when they heard them with distinct realisations of /g/. The fact that this pattern holds across listener groups suggests that the 'distinct standard' is shared at a supra-regional level.

The regional variable: The suffix /əð/

As for the reduction variable, a full model including variant and listener origin as predictors as well as presentation order, word frequency and speaker handedness as fixed effects and individual listener as random effect was fit to the response time data for the critical items of the regional variable. This model was also stepped down, and only the significant factors are reported in the final model summary.

Table 2: Model summary for the regional variable.

Factor	*Estimate*	*Std. Error*	*t value*	*Pr(>|t|)*
(Intercept)	453	23.9	18.9	<2e-16 ***
Listener: Copenhagen	61	26.4	2.3	0.024 *

There was no effect of variant ($p = 0.75$), but, as Table 2 shows, there was an effect of listener origin ($p = 0.024$), with an estimate of 61 (std. error 26.4) for Copenhagen listeners. This means that listeners are equally fast at responding to Aarhus and Copenhagen forms of a word, but Copenhagen listeners are generally slower than Aarhus listeners. Recall that Copenhagener listeners were presented with items in a less familiar accent (the Aarhus tonal contour), which may explain the general increased lag. We return to this latter finding in the analysis of the prosodic pattern below. What is interesting to note here is that the segmental regional variation did not inhibit word recognition for either group. The error rates for words with the regional variable show no effect of listener origin ($p = 0.66$), but there is a general effect of variant ($p < 0.001$), with an error rate of 14% for words with the Aarhus variant as compared to 6% for the Copenhagen variant. In summary, regionally specific segmental variation does not slow down the recognition process, but more mistakes are made, meaning more words with the Aarhus variant are classified as non-words.

Regional prosody

As noted previously, the 32 Copenhagen listeners were presented with stimuli in two different tonal contours: the reduction variable stimuli were presented with a Copenhagen based tonal contour, where the tone on the stressed syllable is low and then followed by a rise to the posttonic syllable. This may conveniently be abbreviated as an L*H pattern (following conventions from the ToBI framework – Beckman, Hirschberg and Shattuck-Hufnagel 2005). The Aarhus variable was presented in the Aarhus based tonal contour, where the tone is high on the stressed syllable and followed by a fall to the posttonic, a pattern that may be abbreviated as H*L. The Aarhus listeners heard all stimuli in an H*L pattern, and, therefore, there is no tonal difference to compare for their responses. But for the Copenhagen listeners, it is possible to examine response times to items that were segmentally possible in either variety and thus mainly differ in tonal contour. I say mainly, because the segmental match between conditions is not as neat as in the two previous analyses. The previous comparisons involved essentially "free variation" in one segment, whereas here the comparison is of disyllabic words with entirely different phonemes, e.g. the response time to [lɔg̊ə] *lukke* ('close' vb.) and [hʌb̥əð̥] *hoppet* ('jumped'). Importantly, however, response times to [lɔɯə] are excluded, because they inhibit processing, and response times to [hʌb̥əd̥] are also excluded, because they would be unfamiliar to Copenhagen listeners in that they would have had relatively little exposure to such forms and they would not ever have used such forms themselves. In this sense, the analysis presented here is an analysis of the effect of prosody, specifically the tonal stress group pattern. The effect of this tonal stress group pattern is interesting not only because a difference was found between Aar-

hus and Copenhagen listeners, but also because these particular tonal stress group patterns have been shown to be important in the identification of speaker regional origin by Danish listeners (Kristiansen, Pharao and Maegaard 2013; Tøndering and Pharao 2020).

A mixed-effects regression model with speaker accent as well as presentation order, word frequency and speaker handedness as fixed effects and individual listener as random effect was fit to the raw response time data for critical items with intervocalic /g/ realised as [g̊] and the suffix realised as [əð̩]. The full model was stepped down and the final model including only the factors that emerged as significant is shown in Table 3.

Table 3: Model summary for regional prosody

	Estimate	*Std. Error*	*t value*	*Pr(>|t|)*
(Intercept)	476	13.7	34.5	< 2e-16 ***
*Tone: L*H*	-59	11.4	-5.2	2.94e-07 ***

The model summary shows that tonal contour is a significant predictor of response times for the Copenhagen participants ($p < 0.001$) with an estimate of -59 (std.error 11.4 when response times to L*H items is compared to response times for H*L contours, meaning that the Copenhagen listeners here were on average 59 milliseconds faster at recognising words in their own L*H accent than in the less familiar H*L accent.

Summarising the results of the lexical decision experiments, we saw that segmental reduction caused inhibition for listeners from both Aarhus and Copenhagen, whereas regional segmental variation did not affect response times for any of the two groups. At the segmental level, then, we may say that deviation from the distinct form has processing costs, whereas regional variation is accepted by the processing system. However, the result from the effect of the difference in tonal contour between the two regional accents for the Copenhagen listeners showed inhibition. This suggests that a more global feature of a regional accent may incur a processing cost to listeners with a different accent. These processing patterns tell us about the implications of standardness at the level of immediate processing, when spoken words are encountered as sound signals and decoded to be matched with lexical items in memory. But how does variation affect the link between variable forms in memory? Is it the case that while (as we have seen) some types of word forms are more easily recognised than others, multiple forms are still stored in memory, and, if so, to what degree are they connected to each other? To study this, we must look at the results from the second phase of our series of lexical decision experiments.

Representation of variable forms: Long-term form priming

In addition to the two blocks of lexical decision tasks that were reported above, the same listeners also participated in a second phase of the experiment, so-called long-term priming tasks with the same variables. This consisted of another series of lexical decision tasks in which the critical items from the first phase served as primes for the critical items in this second phase (as in Sumner and Samuel 2005; 2009). Stimuli for the second blocks were produced by two male speakers, one from Copenhagen and one from Aarhus. Again, speakers were instructed to produce distinct and reduced variants of words with the reduction variable and only the Aarhus speaker produced the stimuli for the regional variable. Speaker gender was changed between the two phases to avoid long-term form priming effects due to voice alone (cf. Goldinger 1996).

In other words, the two new blocks of stimuli were: reduction again and regional again. As before, the order was reversed for half of the participants in each location. Whereas the first two blocks (from phase 1) consisted of 82 trials each, the blocks in phase 2 consisted of 104 trials, as the controls in the long-term form priming condition consisted of words which had not been presented in the first block. It is the difference in response times in phase 2 to these 22 new items compared to the previously encountered items (i.e. the stimuli that constitute the long-term form priming condition), that measures long-term form priming: if response times to these new items are statistically significantly longer than response times to the previously presented items, this means that long-term form priming has occurred. In other words, this would mean that the previous presentation has facilitated the recognition of the repeated word.

Results for long-term form priming

The results of the long-term priming conditions in the lexical decision tasks will reveal to what extent the various forms are stored in the mental lexicon and linked to each other. We begin by looking at the long-term priming results for the reduction variable. Each critical item in these blocks was coded for whether the listener had been presented with an identical target in the first block, a related target or whether this was the first presentation. Using the word /lɔgə/ *lukke* (‘close’ vb.) as an example we can illustrate the different types of stimuli in the following way for the second phase of the experiment: if a participant has heard the form [lɔɥə] in the first phase, this form will be an identical target in the second phase, and the form [lɔg̊ə] will be a related target, because it is a variant form of the same word. This enables us to analyse differences in response times using the last condition, unprimed targets, as the reference level for the regression analysis. To use the previous example as illustration: if only the identical condition is statistically significantly

different from the unprimed condition, this must mean that only reduced word forms can prime reduced word forms. If the related target condition is also statistically significantly different from the unprimed condition, this can be interpreted as a long-term form priming effect of the reduced form [lɔɯə] on the recognition of the distinct form [lɔg̊ə].

Preliminary regression analyses showed a simple main effect of listener origin for both variables, which revealed that the Copenhagen listener group had longer response times than the Aarhus listener group overall. Since the effect did not interact with other factors, the results for each group will be presented separately. The data were again analysed by fitting mixed effects linear regression models to the raw response times with outliers removed (as before, comparison with the results for log transformed response times revealed no significant differences and therefore the raw response times are shown here). Fixed effect factors were: prime-target combination, order of presentation, word frequency of the target and hand preference. Individual participant was included as a random effect.

The reduction variable

For the Copenhagen listeners there was a main effect of prime-target combination only, and both levels emerged as significantly different from the unprimed condition for both the distinct and the reduced variant. The effect is illustrated in Figure 1

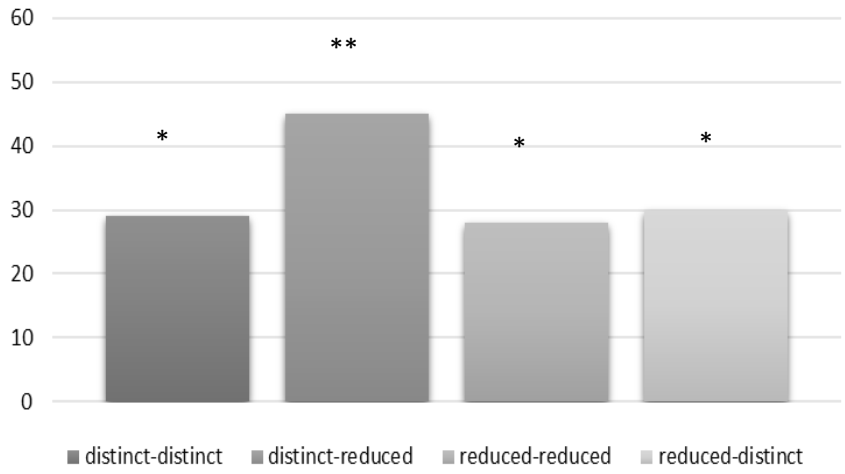

Figure 1: Differences in response times for each of the four priming conditions, intervocalic /g/ – Copenhagen listeners

The unit for the y-axis is milliseconds, and each of the bars represents a priming condition as labelled on the x-axis: the first term in a pair indicates the nature of the prime, whether distinct or reduced, and the second term in the pair indicates the nature of the target, whether distinct or reduced. The height of the bar indicates the mean difference in response times to targets in the given priming condition as compared to the same type of target in the unprimed condition. So the first bar shows that respondents were 29 msec faster to respond to a distinct word form when they had been primed with a distinct form (i.e. repetition priming) as compared to their response time to a distinct word form they had not previously encountered. The second bar shows that listeners were 45 msec faster to respond to a reduced word form when they had been primed with the distinct form of that word as compared to the unprimed condition for the reduced form, etc. All four conditions show a statistically significant difference from the unprimed conditions, meaning that there is repetition priming for both distinct and reduced forms but also that distinct forms can prime reduced forms and reduced forms can prime distinct forms ($p < 0.05$ for all conditions).

The same model was fit to the data from the Aarhus listeners. Again, only prime-target combination emerged as significant. The effect is shown in Figure 2.

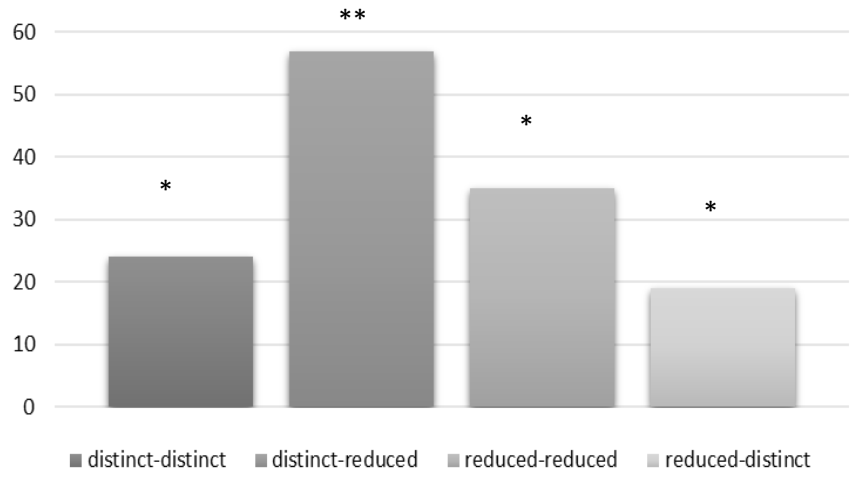

Figure 2: Differences in response times for each of the four priming conditions, intervocalic /g/ – Aarhus listeners

Priming conditions are shown in the same order as in Figure 1 and reveal the same basic pattern as before. Once more, the difference in response times is statistically

significant from the unprimed condition in all four priming conditions; and as for the Copenhagen listeners, the priming effect is strongest for the combination of reduced targets with the distinct primes ($p < 0.05$ in all conditions). The responses to the reduction variable in the long-term priming condition were also analysed for error rates. A mixed-effects logistic regression model was fit to the data with listener origin, prime-target combination and variant as fixed effects and participant as random effect. There was no effect of listener origin ($p = 0.52$) or priming ($p = 0.63$), but only a slight effect of variant as before ($p = 0.031$), with an error rate of 4% for the words with reduced variants, compared to just 1% for words with the distinct variant. This effect of variant also holds for the subset of unprimed stimuli, and the overall improvement in accuracy can therefore not be attributed to priming itself, but is perhaps due to a more general learning effect of participating in the experiment. This means that Aarhus and Copenhagen listeners were equally good at correctly identifying real words, even though Copenhagen listeners were slower. So priming did not improve accuracy on its own, even though the error rate for words with reduced variants was somewhat reduced compared to the results from the first phase of the experiment.

The regional segmental variable

Next, we look at the long-term priming results for the regional variable. As before, response times were generally longer for the Copenhagen listeners than for the Aarhus listeners, and therefore the regression models were fit to the two datasets separately. Again, targets were coded for priming condition, so that using the unprimed condition as the reference level for each variant would reveal whether response times in the individual priming conditions are statistically significantly different from those in the unprimed conditions. The same mixed-effects models as for the reduction variable were fit to raw response times with outliers removed.

For the Copenhagen listeners, none of the factors emerged as significant, including the priming conditions. This means that for the Copenhagen listeners, not even repetition priming occurred for the segmentally Copenhagen forms ($p = 0.24$), and therefore the differences in response times for the priming conditions are not shown. This result is surprising, but recall that stimuli for the regional variable were always produced by speakers from Aarhus, and thus realised with the regional H*L stress group pattern. The results suggest that the stress group tonal pattern alone blocks the possibility for priming for listeners who do not produce this pattern themselves.

The same model was fit to the data from the Aarhus listeners. Here, the priming condition emerged as statistically significant. The effect for the Aarhus listeners is shown in Figure 3. Note that the only statistically significant effect is the repetition priming for the Aarhus form. In other words, for the Aarhus listeners, only the Aar-

hus form can prime the Aarhus form and neither variant primes the Copenhagen form. This result is somewhat surprising, as in this case listeners are familiar with the intonation pattern of the stimuli and use it themselves.

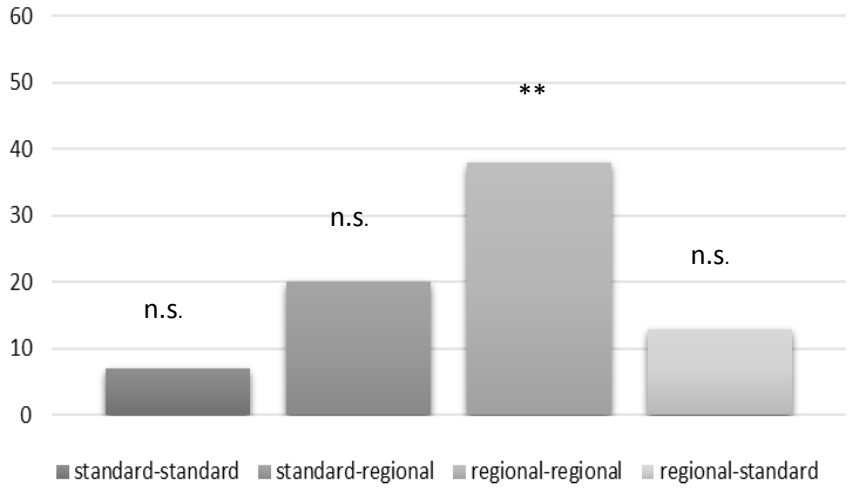

Figure 3: Differences in response times for each of the four priming conditions, /əð/ – Aarhus listeners

As for the error rates for the regional variable, there was no effect of listener origin (p = 0.25) or variant (p = 0.48) but there was a slight yet statistically significant effect of priming on the error rates, with words with the Copenhagen variant having an error rate of only 1% when they had been primed with the Copenhagen variant, as compared to the unprimed words with Copenhagen variants, which had an error rate of 3%.

Taken together, the results for both variables suggest that intonation patterns play a significant role in processing and particularly storage of spoken words. Recall that the stimuli for the reduction variable were always presented to listeners in the tonal pattern they were familiar with and use themselves. This is possible because the distinct and reduced variants of /g/ occur in both varieties of Danish. In other words, both variants are associated with both stress group patterns in speech production. This is not the case for the regional variable, since one of the segmental variants only occurs with Aarhus intonation. The results suggest that there is a strong association between the Aarhus segmental variant and the Aarhus stress group pattern, and that such an association is needed in order for priming to obtain.

This association is also there for the reduction variable, as shown by the long-term priming effects for listeners in both groups; and since both distinct and reduced variants may occur with both types of tonal contour, links appear to exist between both types of segmental forms, although the priming effect of distinct forms on reduced forms was stronger overall, again suggesting a processing preference for distinct forms.

To examine the effect of the tonal pattern on lexical storage in more detail, we may again confine ourselves to the data from Copenhagen participants, who were exposed to both their own L*H contour and the H*L contour. By looking at what we have been calling segmentally Copenhagen forms only, we can analyse the response times in terms of tonal contour alone. It is only possible to examine repetition priming here because of course the stimuli in the two different contours were not related in any way segmentally. When the same model as before is fit to this subset and stepped down, we only find a significant long-term priming effect for the L*H items ($p < 0.01$), never for H*L items ($p = 0.3$). This is to be expected from the results presented above, and this additional analysis simply supports the interpretation that the prosody must be not only familiar to the listener in order for priming to obtain, but the prosody should also be associated with the segmental variant. In other words, less familiar prosodic patterns appear to block the formation between variable lexical items in memory.

STANDARD DANISH IN THE MINDS OF LISTENERS

For intervocalic /g/, it was found that reduced forms inhibited word recognition for all listeners, but also that both distinct and reduced forms could induce priming, indicating that both forms are stored in the mental lexicon. However, the priming effect was stronger for distinct forms. Taken together, this suggests that distinct forms are more strongly encoded in the mental lexicon than their reduced counterparts. For the regional variable, the suffix /əð/, no difference was found in terms of immediate processing of the segmental content, since Copenhagen and Aarhus forms were recognised at the same speed for both groups of listeners. However, regional prosodic variation affected response times for the Copenhagen listeners who were less familiar with the tonal contour in the regional items and both segmental and prosodic regional variation severely influenced long-term priming. There was no long-term priming for the Copenhagen listeners at all for the regional variable and only repetition priming for the Aarhus variant for the Aarhus listeners, who use the variant themselves in combination with the stress group pattern that the regional variable stimuli contained. As shown by the relatively low error rates, it is of course not the case that words with phonetic features typical of Aarhus cannot be

recognised by listeners from Copenhagen. But it appears that the tonal contour incurs a cost in immediate processing and also blocks the formation of links between variable forms in memory. It is not possible to tell from this dataset whether the effect of the tonal contour is due to it being less familiar to the listeners, as we do not have comparable data from the Aarhus group. A follow-up experiment is needed to gain further insight into the role of exposure for how prosodic differences influence lexical processing.

What do these findings tell us about the status of standard Danish in terms of the mental processing and representation of variable forms? If we look at the results together, they would appear to indicate that overt attitudes towards speech variation shape lexical processing. In terms of distinctness, we only have knowledge of overtly offered attitudes in public debates about proper ways of speaking Danish. People express difficulty in understanding what they experience as speech with reduced forms, and not a few find such speech aesthetically unpleasing. The results show that processing of reduced forms is in fact more difficult in the sense that it takes longer and more often leads to errors. We cannot know from the data currently available whether this processing cost is caused by some phonetic property of reduced forms, e.g. less variability in terms of the stream of acoustic cues in the reduced word forms, that makes them more difficult to recognise, or whether the negative attitudes towards the reduced forms cause a delay in the processing system. A more direct investigation of the link between attitudes and word recognition is therefore an important avenue for future research. But the findings for the reduction variable do show that exposure alone cannot explain the path of processing, since reduction of intervocalic /g/ is more common in running speech than non-reduction. Word recognition cannot therefore simply be a matter of matching the most frequently occurring phonetic pattern to word forms stored in memory. I would suggest that the results are best understood in terms of a privileged role of the less frequent distinct forms in the memory of speakers (as also suggested in the model proposed in Sumner et al. 2014). As such these distinct forms can be viewed as the standard forms in the minds of the language users as well as in the prescriptive norm advocated for public speaking (although these are mostly notable in complaints about "sloppy speech" in news reports and films, or in guidelines for journalists working in broadcast media). In other words, the results for the reduction variable support an interpretation where an explicit evaluation that shows preference for distinct forms matches the "preference" for the implicit and automatic processing system in the minds of the language users that allows them to recognise and therefore understand spoken word forms. However, as will be discussed more extensively below, it is important to keep in mind that this finding is based on results obtained for words presented in isolation, a task that is not straightforwardly related to the recognition of words in the speech stream of everyday conversation.

The findings for regional variation are more complex. We know from the language attitude studies reviewed above that Danes overtly prefer their own local variety, but covertly prefer Copenhagen-based speech. Looking at the findings for segmental variation alone, it would appear that neither of these preferences play a role in processing, as word forms with segmental variants particular to the Aarhus region [əd] were recognised as fast as word forms with segmental variants that also occur in Copenhagen-based speech [əð̩]. This might suggest the existence of regional standards that allow for socially stratified variation while still privileging distinct forms. But what about the finding that Copenhagen listeners took longer to process words with Aarhus tonal patterns, both as compared to the Aarhus listeners and compared to their own performance with words with the Copenhagen tonal pattern? Add to this the complete lack of priming for Aarhus forms in the Copenhagen listeners, and the findings certainly show that regional varieties do not enjoy the same status (in processing terms) across speech communities in Denmark. I suggest that these findings taken together with those from the Aarhus listeners do argue in favour of the existence of regional standard Danishes.

The findings from the reduction variable have already called into question the role of exposure in the processing of variable forms. The same can be said for the findings concerning the regional segmental variable, although with a different effect: while the realization [əd̥] does occur in the speech of Aarhus speakers, it is not the dominant form (Jensen and Maegaard 2012). Yet words ending in [əd̥] are recognised as quickly as the same word ending in [əð̩] by both groups of listeners. If exposure played a role in recognition, there ought to have been a difference for both groups of listeners. Additionally, we might have expected this effect to be additive for the Copenhagen listeners. Instead, regionally based segmental variation seems to be tolerated by the processing system.

The finding for the tonal patterns might suggest a role for exposure and/or familiarity, but we may also interpret this with reference to the overtly offered language attitudes found in the label ranking tasks conducted by Kristiansen: all participants, including those from Copenhagen and those from Odder (a town just to the south of Aarhus), indicated a preference for their own local way of speaking. As previous studies have shown, the tonal stress group pattern plays a significant role in the identification of a speaker's regional origin. It seems plausible, then, that the delay in recognition associated with the Aarhus tonal pattern for Copenhagen listeners may be due to the additional task of identifying the speaker as being from a different region than themselves. Once this identification is accomplished, segmental variants are accepted without processing costs. In this sense, language users appear to accept the existence of different regional standards. But upon encountering reduced variants, which are associated with any and all regional varieties, the same language users are challenged in their recognition of spoken words.

This series of studies is only a first step in understanding the role of standards of spoken Danish in everyday communication. As is so often the case with psycholinguistic studies, the interpretations presented here and their association with talk in interaction outside an experimental setting rest on the assumption that language users do not have separate processing systems for different contexts, one for the lab and one for the home, as it were. While this seems to me to be incontrovertible, we must still acknowledge that the recognition of spoken words in isolation is a different task than recognising and decoding words in running speech, as most contextual cues are removed in the lexical decision task. Findings from other languages (e.g. Tucker 2011), as well as a pilot study using the same Danish stimuli as were used here for the reduction variable but presenting them in utterances, suggest that context does not ease the processing of reduced forms: words with distinct segmental variants are still recognised more quickly. Further experiments studying the processing of words in context would help shed light on the preference for distinct forms found here. It would also be useful to study in greater detail the attitudes towards reduced variants held by Danes. While the interpretations presented here suggest that overtly offered attitudes are most closely aligned with the mechanisms of the processing system, studies involving the evaluation of (speakers using) word forms with reduced variants would be able to shed further light on how the difference observed here between the two types of segmental variation is shaped by covertly held attitudes towards speech variation. In spite of these shortcomings, I hope to have shown how the use of psycholinguistic tools can provide a more nuanced understanding of the standard.

There is a wealth of studies that take an evaluative approach to the question of the standard language and how to define it in a given context. In most cases, overtly codified prescriptive norms refer only to written language, e.g. official standard dictionaries or particular guides on usage and style in writing. Norms for spoken language also exist and can be studied in terms of shared ideologies within speech communities. Such community level standards can be passed on from generation to generation along with the variable speech patterns that the standard proscribes against. In other words, a language user learns what counts as the best or even correct way of speaking is, whether this is in terms of syntax, morphology or accent, but does not necessarily adhere to this standard, certainly not across all social contexts. In other words, while attitudes towards different ways of speaking may influence language change, speakers can still choose not to follow the community norm when they talk. This is why including data on how variable word forms are processed is interesting when we want to understand the standard. The mechanisms of the processing system and in particular the speech perception and comprehension systems are automatic and very difficult if not impossible for the language user to control (e.g. we cannot decide not to hear words, and in many so-called slips of the

ear, words are misinterpreted as other words, not as nonsense). So the listener cannot choose to find some forms easier to recognise than others. Therefore, preference for one form over another or indeed as here a lack of any clear preference can be taken as evidence for what counts as the standard spoken language in the mind of the individual listener. It is not a given that words with intervocalic stops should be more easy to recognise than words with intervocalic approximants, but the patterns in response times found here shows that this is so for Danish listeners. In this sense, the distinct forms are the standard forms in spoken Danish – not because 'expert orators' have decided that it should be so or because generations of parents have taught their children this, but because distinct forms make word recognition easier. It is important to keep in mind that the findings presented here are based on words presented in isolation. In such a situation, the listeners are left without contextual cues as to the meaning of the words and non-words that are presented to them. The recognition of the meaning is crucial to the task they are performing, and it is possible that the lack of contextual cues lead to a greater reliance on canonical forms. That is to say, we cannot conclude from the results reported here that distinct forms will *always* be easier to recognise than reduced forms: this may depend on the context and it may certainly also be related to the evaluations of speech containing distinct forms, although this remains to be investigated directly. The results for the regional variable reported here suggest that we should not expect to find a direct link between processing and implicit evaluation. Instead, it may well be that there is a national ideology in which one particular accent is the preferred or standard accent (and in Denmark that would be the Copenhagen-based accent), but that this does not affect the processing system. In other words, language users may be able to handle a variety of regional standards, even if the community they belong to only officially have one standard language.

* The data used in this chapter was collected by Mia Ridder Malmstedt for her MA thesis in Linguistics under my supervision. The analyses reported here are all new.

REFERENCES

Beckman, M., J. Hirschberg and S. Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (ed.) *Prosodic models and transcription: Towards prosodic typology*. Oxford University Press: UK. 9–54.

Ernestus, M., R. H. Bayeen and R. Schreuder. 2002. The recognition of reduced word forms. *Brain and Language* 81: 27–41.

Floccia, C., J. Goslin, F. Girard and G. Konopczynski. 2006. Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance* 5: 1276–1293.

Goldinger, S. D. 1996. Words and Voices: Episodic Traces in Spoken Word Identification and Recognition Memory. *Journal of Experimental Psychology Learning Memory and Cognition* 22, 5: 1166–1183.

Grondlaers, S. and T. Kristiansen. 2013. On the need to access deep evaluations when searching for the motor of standard language change. In T. Kristiansen and S. Grondelaers (eds.) *Language (De)standardisation in Late Modern Europe: Experimental Studies*. Oslo: Novus. 9–52.

Jensen, T. J. and M. Maegaard. 2012. Past participles in Jutland Danish: A real time study of regionalization and standardization. *Nordic Journal of Linguistics* 35, 2: 169–195.

Kristiansen, T. 2009. The macro-level social meanings of late-modern Danish accents. *Acta Linguistica Hafniensia* 41: 167–192.

Kristiansen, T., N. Pharao and M. Maegaard. 2013. Controlled manipulation of intonational difference: An experimental study of intonation patterns as the basis for language-ideological constructs of geographical provenance and linguistic standardness in young Danes. In T. Kristiansen and S. Grondelaers (eds.) *Language Destandardisation in Late Modern Europe: Experimental Studies*. Oslo: Novus. 355–374.

Maegaard, M., T. J. Jensen, T. Kristiansen and J. N. Jørgensen. 2013. Diffusion of language change: Accommodation to a moving target. *Journal of Sociolinguistics* 17, 1: 3–36.

Pharao, N. 2011. Plosive reduction at the group level and in the individual speaker. *Proceedings of the 17th International Congress of Phonetic Sciences.* Hong Kong, China. 1590–1593.

Sumner, M. and A. G. Samuel. 2005. Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language* 53, 2: 322–338.

Sumner, M. and A. G. Samuel. 2009. The role of experience in the processing of cross-dialectal variation. *Journal of Memory and Language* 60: 487–501.

Sumner, M., S. K. Kim, E. King and K. McGowan. 2014. The socially-weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology* 4: 1–13.

Tucker, B. V. 2011. The effect of reduction on the processing of flaps and /g/ in isolated words. *Journal of Phonetic* 39: 312–318.

Tøndering, J and N. Pharao. 2020. Recognising regional varieties of Danish. *Linguistics Vanguard* s1 https://www.degruyter.com/view/j/lingvan.2020.6.issue-s1/lingvan-2019-0020/lingvan-2019-0020.xml (March 1, 2022).

Veenker, T.J.G. 2018. The Zep Experiment Control Application (version 2.2.2) [software]. Beexy Behavioral Experiment Software. Retrieved from www.beexy.org/zep/.