Sound change, priming, salience

Producing and perceiving variation in Liverpool English

Marten Juskan

Draft
of 9th October 2018, 10:13

./langsci/graphics/la

Language Variation

Editors: John Nerbonne, Dirk Geeraerts

In this series:

1. Côté, Marie-Hélène, Remco Knooihuizen and John Nerbonne (eds.). The future of dialects.

2. Schäfer, Lea. Sprachliche Imitation: Jiddisch in der deutschsprachigen Literatur (18.–20. Jahrhundert).

3. Juskan, Marten. Sound change, priming, salience: Producing and perceiving variation in Liverpool English.

Sound change, priming, salience

Producing and perceiving variation in Liverpool English

Marten Juskan

Marten Juskan. 2018. *Sound change, priming, salience*: *Producing and perceiving variation in Liverpool English* (Language Variation 3). Berlin: Language Science Press.

./langsci/graphics/storagelogo.pdf

To Daniela

Contents

Draft of 9th October 2018, 10:13

12 Discussion (Perception)

12.1 Priming and salience

12.1.1 Salient vs. non-salient variables

The main hypothesis of this study is that priming effects in perception experiments depend on the salience of the variable that is investigated. This is largely corroborated. No priming effect could be identified for HAPPY, the non-salient vocalic variable: priming condition did not surface as a significant predictor. This suggests that subjects were not influenced by social information relating to the speaker's regional origin. Rather, they perceived the stimuli in pretty much the same way, regardless of whether they had been told the speaker was from Manchester or Liverpool. In addition, answers were comparatively evenly distributed across all four synthesised tokens, the preference for tokens 2 and 3 was only weak. It seems thus that subjects were not too sure (or unanimous) about which token best matched the stimulus heard in the sentence.

Essentially the same remarks can be made about the responses to velar nasal plus stimuli. Overall, the distribution of answers looks very similar to those collected for HAPPY. However, the dominance of tokens 2 and 3 answers has clearly increased compared to HAPPY responses: These two tokens together now account for around 75% of answers (as opposed to about 65% for HAPPY). In the mixed-effects ordinal regression model, which also took into account random variation due to individual characteristics of the participants, there was no significant difference between priming conditions. All the same it should be noted that even in this mixed-effects regression, the p-value for prime indicated a statistical trend. Both in terms of overall distribution and statistical significance it could therefore be said that the /ŋ(g)/ data occupy a sort of middle ground, which ties in quite nicely with the production results, where this variable also ended up in front of HAPPY but behind both NURSE and /k/, as far as salience is concerned.

Towards the other end of the scale we have NURSE and lenition of /k/. When confronted with NURSE stimuli, people have a very clear preference: Token 2, the synthesised vowel that is closest to the one actually heard in the sentence, is chosen in around 2 out of 3 cases. The other options only account for a rather

small minority of cases. In the case of NURSE stimuli, subjects were thus much better at identifying the best match, and responses were much more uniform. In addition, participants were more susceptible to the priming manipulation for NURSE than for both happy and velar nasal plus. Subjects perceived NURSE stimuli differently in the two priming conditions, and this difference was found to be significant in the mixed-effects regression, which identified prime as a main effect (but also showed that the effect was mostly driven by sentence-medial stimuli).

Results for lenition of /k/ are similar to those for NURSE in the same way that happy results are comparable to the ones found for /ŋ(g)/. We again find token 2 to be the dominant response, but the /k/ results differ from the NURSE data in that the dominance of the objectively most accurate answer option is even more pronounced here: Token 2 alone has a share of almost 75% of all answers (just above 65% in the case of NURSE). Subjects again behaved differently, depending on which city/accent they had been primed for: The priming effect was found to be significant in the regression model, provided the interaction of prime and social class was removed first. Judged against the background of the very strong preference for one particular token, the priming effect for lenition stimuli seems to be even more statistically robust than it was for NURSE.

12.1.2 Degree of priming and accuracy

So far, two criteria have been identified that can be used to place the four variables investigated on a scale: Objective accuracy of responses and degree of the priming effect. With respect to priming, we would have happy at the lower end of the scale, because this variable did not produce any priming effect at all. Next would be velar nasal plus, where a priming effect might be suspected when looking at the raw data, but becomes non-significant in the mixed-effects ordinal regression. This would be followed by NURSE, which produced a clear priming effect that was statistically robust. At the upper end of the scale, finally, we find /k/-lenition, the variable where a statistically robust priming effect could be found even though, overall, subjects were heavily focused on the objectively correct answer token. When we look at accuracy, the same picture emerges: Matching the synthesised tokens to the stimuli seems to have been most difficult for happy, where responses are very diverse, and the most accurate token accounts for less than 40% of answers. In the /ŋ(g)/ data, this percentage is slightly higher, and, more importantly, the share of token 4 (the one that is most unlike the vowel actually heard in the sentence) drops considerably. Perception of NURSE stimuli was then considerably more accurate ('correct' percept in 2 out of 3 cases) and in the case of lenition, finally, participants chose the objectively most similar option almost 75% of the time.

It is possible that the *task* participants had to perform was not always equally difficult, either because some of the four variables were *intrinsically* more 'difficult' than others, or for reasons of experimental design and stimulus creation. I do not see any compelling evidence for this argument, though. For the two vowels in particular the extremes of the answer scale were very comparable: A rather central vowel on the lower end, and a fronted (and raised in the case of happy vs. lowered in the case of NURSE) one on the upper end. The parameters used in synthesis were also quite similar (cf. Table 10.2), so there should have been roughly equal phonetic distance between answer tokens of both variables. For the two consonants this kind of equivalence is harder to achieve, but remember that here, too, a comparable feature (proportion of frication/aspiration) was manipulated in both cases (cf. §5.4.1 and §10.1.3). In addition, the resulting tokens were not only checked auditorily by the author, but also subjected to the scrutiny of other linguists during a pilot test which did not reveal any problems with the stimuli or the synthesised answer tokens. We know from work about folk linguistics and perceptual dialectology that experts' judgements do not necessarily have to coincide with those of laypersons (Preston 1999; Niedzielski & Preston 2000). But while it cannot be ruled out that perceivers found the tokens of one variable to be less distinct than those of another, there is no evidence to actively support this idea (in the post-experiment comments of subjects, for instance).

I will therefore assume that there is another explanation, and one which is capable of explaining both the differences in terms of accuracy and in the priming effect. It is, after all, quite striking that these two criteria result in exactly the same ordering of variables. This parallelism suggests a common source or characteristic, and I believe this characteristic is social salience. If a variable is socially salient, i.e. more informative in social terms, this will have consequences for the perception, and, more importantly, the storage in long-term memory of exemplars pertaining to this variable. As explained earlier in this study, not everything we experience is memorised (cf. §4.2.3). In fact, we cannot even actively *process* every little detail present in the visual or acoustic input, let alone *store* all of it. Rather, we pay attention to certain things and not to others, and only the information that passes through this filter enters into long-term memory.

12.1.3 Salience as likelihood of remembrance

Whatever definition of salience one adheres to, the one common feature everyone usually can agree on is that salient features 'stick out' (cf. §4.1.1). Another way of putting this is to say that a salient feature attracts attention. Chapter 4 explained that, in an exemplar framework, salience should therefore have an im-

pact on memory structure. Salient features (as well as the words and phrases containing them) will be more often remembered than non- or less salient ones. The result is that exemplar clouds of salient variables will be more detailed because they contain a greater number of slightly different variants due to the fact that realisations of this variable have a higher likelihood of being remembered. For the same reason (likelihood of remembrance) the most common realisations will also be more entrenched because these exemplars get strengthened more often by newly remembered, largely identical input, than the traces of non-salient variables do.

If we assume that features are at least in part salient because they are socially informative it is also conceivable that salience has an impact on indexation, in the sense that less additional information will be remembered if the variable is not salient. When a perceiver does not believe a feature to be socially diagnostic they will pay less attention to it. This will either mean that social information is remembered, but fades more quickly than for socially more meaningful exemplars, or the exemplar is not indexed with certain kinds of social information in the first place. After all, if some sort of acoustic input is not (thought to be) linked to a specific social group, why store information about it?

The priming differences found in my data are hard to account for in a non-episodic framework. After all, if, say, some general sort of social filter (cf. Niedzielski 1999) interfered with subjects' perception, why would this apply more to some variables than to others? If the prime 'Liverpool' triggererd expectations based on knowledge about Scouse segment realisation rules, then why do we find differences between the variables, and especially differences in degree. Presumably, a rule is either known (in which case it should trigger a priming effect) or unknown (in which case the prime should be ineffective), but not 'slightly more known' than the rule for another variable. It is not immediately obvious how a non-episodic explanation would account for the fact that not all parts of the input (i.e. the previous experience with Scouse) seem to enjoy equal prominence in long-term memory.

If, however, priming builds on remembered exemplar clouds and if these clouds are variably structured, differences in the priming effect are only to be expected. If social information is used to prime for a socially non-salient variable, either no exemplars are activated at all (because they are not indexed for this kind of information), or the activation will be comparatively weak (because there are only a limited number of relevant exemplars that have not faded yet). As a result the priming effect will be weak or even non-existent. In the case of a highly salient variable, on the other hand, priming can rely on a large cloud of exem-

plars which activate both easily and strongly since they have a higher baseline of activation to start with, given that they are strengthened rather frequently. The consequence is a statistically robust and comparatively strong priming effect. This is precisely what was observed: No or weak priming effects for the non-salient variables happʏ and velar nasal plus, stronger, more robust priming effects for the two stereotypes ɴᴜʀsᴇ and lenition of /k/.

Accuracy and the lack thereof is a result of the size of the exemplar cloud, which, as has been explained above, is ultimately a product of salience, too. If subjects are not used to paying attention to a variable, they will only have a small number of memory traces to compare the input to, and it will be more difficult to match the synthesised tokens to the stimulus because the scale available in memory is not very fine-grained. When there *is* a very detailed memory cloud the likelihood that the input will activate a similar exemplar is higher, in which case the subject will feel more confident in making a choice in the experiment. In other words, if remembered exemplars are few and far between, there is a higher probability for the stimulus that needs to be classified to have no exact match in subjects' memory, but instead to be equally distant from a number of different exemplars – all of which are then an acceptable choice – and subjects may then categorise the same input as belonging to different categories on different occasions, without a clear preference for one category in particular. In a dense memory cloud, however, chances are that the stimulus will correspond very well to one remembered exemplar in particular, which means that it will be classified as belonging to that category in the majority of cases. The data collected for this study exhibit just the distributions that are to be expected if one embraces the explanation given above: From the least salient variable happʏ over velar nasal plus and ɴᴜʀsᴇ to /k/ lenition (the most salient one) answers are less and less equally distributed across the four available tokens because the percentage of the objectively correct token increases steadily.

My data furthermore suggest that salience in exemplar priming experiments is not a categorical matter in the sense that a variable is either salient or non-salient, and therefore generates a priming effect or does not. Rather, the word *scale* was used deliberately when describing the ordering of the four test variables earlier. While the non-salient variables happʏ and /ŋ(g)/, and the salient ones ɴᴜʀsᴇ and /k/-lenition, respectively, do to a certain extent behave as a group, it has also been outlined that there is evidence for a continuum: Perception of velar nasal plus is slightly more primeable (statistical trend) and accurate than it is for happʏ, and perception of /k/ seems slightly easier to prime than ɴᴜʀsᴇ, given that, in the former case, the effect is robust *despite* an extremely strong preference

for token 2. Further research on a larger set of variables is needed to see if this pattern can be replicated, but the fact that the same ordering also emerges from the production data (cf. §9.5) lends support to this interpretation. Furthermore, if we posit – as I have done above – that salience translates into attention paid to a variable in perception and that its effect can be operationalised as the likelihood of remembrance of an exemplar (or the likelihood of indexation with a certain type of information), then there is no reason *not* to assume that this phenomenon is gradual in nature. Since a likelihood can assume a theoretically infinite number of concrete values, there can also be an infinite number of exemplar clouds that differ in terms of size/detail and the number of exemplars that are indexed for a specific category.

12.2 Social factors

12.2.1 Social class

I would argue that the gradualness of salience also (indirectly) shows up in the social characteristics of the *perceiver* that play a role in priming. Hay, Nolan, et al. 2006 and Hay & Drager 2010, for instance, found an interaction of priming condition and social class of the participant in their experiments. Only subjects from higher social classes showed a priming effect, whereas those with lower socioeconomic status did not. This was true irrespective of whether priming was achieved with the help of explicit regional labels on the answer sheets or through the presence of stuffed toys that invoked the same concepts (cf. Hay & Drager 2010: 878). The authors explain this effect with the "amount of exposure that New Zealanders from different socioeconomic backgrounds would have to the speech of Australians" (Hay & Drager 2010: 878). They hypothesise (probably rightly so) that New Zealanders from higher social classes are "more able to travel to Australia" and therefore have "more stored exemplars indexed with 'Australian'" which can be activated by a prime.

Amount of exposure is the only explanation Hay & Drager give for their social class interaction. They do not consider the option at all that there might not be the same degree of sensitivity (or attention) to a socially meaningful variable in the different social groups. This is somewhat surprising given that the authors do make reference to this idea when it comes to the influence of perceivers' gender, arguing that "females may be more aware of the relationship between variability in speech and social characteristics and may therefore index their exemplars with a larger amount of social detail and/or place more weight on this social detail"

(Hay & Drager 2010: 884). To be fair, they then go on to discard this explanation (which was first voiced by Drager 2005), because they believe it does not explain all their gender-related results.

The crucial result of the present study in this respect is now that a significant interaction of priming condition and social class of the participant could only be identified for one of the four variables, lenition of /k/. It does make sense to assume that middle-class subjects are more mobile than working-class participants, and might therefore have more exemplars that are indexed with 'Liverpool' and can be activated by social priming. This interpretation works fine for /k/, but it fails to explain why it *only* works for /k/. Why is there no interaction of prime and social class in the NURSE data? It seems very unlikely indeed that middle-class speakers are more often exposed to Scouse variants of /k/ than their working-class counterparts, but have the same amount of experience with Liverpool NURSE variants. As an alternative explanation, I would like to argue once more for awareness, or, in the terminology used above, attention. It has been shown that NURSE – though undoubtedly a salient variable – still seems to be somewhat less salient than lenition of /k/, but this alone would not explain why there is a difference between social classes for the latter, but not for the former. Rather, it would have to be the case that the *difference* in salience between working-class and middle-class listeners is greater for /k/-lenition than it is for NURSE.

This is post hoc argumentation and in need of further support. All the same, this claim does make sense given that lenition of /k/ is arguably not only the most salient feature of Scouse, but also possibly the most stereotyped one which, in addition, is often very negatively evaluated even *within* Liverpool (cf. §8.2.3). It does not seem too far-fetched that middle-class attitudes would be particularly 'extreme' for a feature which is even looked down on by some of its users. While this interpretation of the nearly-significant prime X social class interaction for /k/ ties in quite nicely with the explanation given above for the results pertaining to accuracy and robustness of the priming effect, I would like to stress once more that the sample of this study is heavily skewed towards middle-class participants and that the claims just made about social class differences can only be rather speculative in nature.

12.2.2 Gender

There is even less evidence for the impact of other social factors which are usually of interest in a sociolinguistic study. The perhaps most surprising result is that, in the present data set, gender of subject does not surface as a significant fixed effect in *any* regression model, irrespective of whether the dependent vari-

able's salience is high or low. Previous research has produced heterogeneous evidence in this respect. Niedzielski (1999: 69 and 79–80) found in her study in Detroit that "there was essentially no difference between what male and female respondents selected in either the 'Canadian' group or the 'Michigan' group", despite the fact that only female subjects overtly commented on stereotypical features of Canadian English (but Niedzielski also provides some evidence why she still believes Detroit women and men hold essentially the same stereotypes about Canadian English). Hay and colleagues, on the other hand, consistently did find a gender effect when they replicated Niedzielski's experiment in New Zealand. Their results showed that only female participants behaved as expected, hearing Australian vowels when being primed for Australia, and perceiving more New Zealand vowels when the concept 'New Zealand' had been invoked. Men were also influenced by the prime, but the priming effect was in the opposite direction: Male subjects actually heard more *New Zealand* vowels when they had been led to expect a speaker from Australia (Hay, Nolan, et al. 2006; Hay & Drager 2010). The absence of a gender effect in the present study might be due to the fact that female and male participants share the same stereotypes about Liverpool (and possibly Manchester) English, as Niedzielski has argued, but there might actually be more to this. Both the lack of a gender difference and the direction of the priming effect in this study will be further discussed in §12.4.

12.2.3 Age

Another social factor besides class (which is only relevant for /k/) that comes at least close to statistical significance is age of participant. Both in the data for happʏ and for velar nasal plus there was a statistical trend for older subjects to choose lower number tokens (more Mancunian/standard) more often than younger participants did. As has been noted earlier, this could be explained by language change in the case of happʏ-tensing: The peripheral [i] is now the norm in standard British English, but up until the beginning of the 80s traditional RP speakers would have a lax [ɪ] realisation for this vowel (cf. Harrington 2006). Younger participants have not experienced this change and only know [i] as the standard pronunciation (although they could encounter [ɪ] when listening to very conservative speakers). It is nonetheless dubious whether this is enough to claim that the [ɪ]-[i] distinction in happʏ is more salient for older subjects or that younger speakers are more likely to expect [i] generally. After all, large parts of (northern) England still have a lax happʏ vowel, a variant which is even actively exploited as an identity marker by some younger speakers in Nottingham (cf. Flynn 2010), and, apparently, also in Liverpool (cf. §9.1.1).

In addition, slightly higher salience in the group of older participants would only explain why the distance between priming conditions seems to get a bit larger, but not why older subjects choose lower number tokens more often *across the board*, i.e. irrespective of priming condition. Hay & Drager (2010: 878–879) also found an effect of age in their data: Younger participants were more likely to perceive a tense [i] instead of more central variants. The authors explain this by suggesting that participants at least in part process the input using their own production as a point of reference. Since the younger participants in their study are thought to have more central /ɪ/ realisations than the speaker who produced the stimuli, they are therefore inclined to perceive these stimuli as more peripheral because they *are* relative to their own production. If we want to apply this explanation to the results of the present study, this would mean that older subjects would have to have tenser realisations of /ɪ/ than the younger participants, which would make the stimuli sound more central to the former than to the latter. While this is possible, the scenario does not seem likely. Participants came from all over Britain and as far as I am aware there is no evidence that happʏ is becoming more central in *all* younger speakers across the country. I cannot verify this at present because no production data were collected from participants. Further research would be needed to shed light on this matter.

The slight age effect in the responses to /ŋ(g)/ stimuli is just as difficult to explain. At least for Liverpool speakers, (non-)salience of this feature seems to be stable across age groups (cf. §9.3.3) I have no way of knowing whether this is true for the subjects from the rest of the country as well, but I am at least not aware of any evidence that would suggest anything to the contrary. Do older participants' realisations of <ng> contain voiced velar plosives more often than younger subjects', then? Since most participants are not from the area where velar nasal plus is commonly found this seems rather far-fetched, especially so when considering that the proportion of subjects who come from Trudgill's 1999 velar nasal plus area is actually higher among those aged 35 or younger than among older participants. I do not have a good explanation for the age effects in both happʏ and velar nasal plus results, but it should be noted that there is no call for over-interpretation anyway: (a) For velar nasal plus it is only a statistical trend, in the case of happʏ, age is only a significant predictor in a model that was eventually not even retained, and (b) the sample is heavily skewed towards participants in their twenties. The few older subjects I do have might not actually be representative of the group they have been assumed to be representative of, and their responses might be considerably overlaid by idiosyncracies, due to the small number of participants the answers can be averaged across. A much more

balanced sample (and possibly relevant production data from participants) would be necessary for a more detailed analysis.

12.2.4 Geographical distance

When we look at where participants live (the other social predictor which did not quite reach statistical significance) the picture is also somewhat unclear, albeit for different reasons. Geographical distance from Liverpool is a near-significant fixed effect in the regression models estimating the responses to happʏ and /k/-lenition stimuli. In both cases there is a trend for subjects who live further away from Liverpool to choose higher number tokens somewhat more frequently. That is to say that the further a participant lives from Liverpool, the more likely they are to perceive a more Liverpool-like token. In §11.4 it has been suggested that familiarity might be a possible explanation for this effect. If we assume that people who live close to the Liverpool area are more familiar with the accent of this city (and I believe this makes sense) it could be the case that these subjects are less likely to perceive the stimuli as one of the Liverpool variants that are represented by tokens 3 and 4, simply because they know rather well what these realisations actually sound like (i.e. they have more stored exemplars of these variants), and, as a consequence, feel more secure in deciding that this is not what was presented in the stimuli. People who live far away from Liverpool and have little personal experience with Scouse, on the other hand, might be more tempted to select tokens 3 and 4 from time to time, feeling less confident to rule them out and possibly assuming that these answer options must be there 'for a reason'.

This interpretation might seem rather speculative, but it would at least explain why there was no interaction of distance and prime: We are not seeing an effect only in the 'Liverpool' group, but actually in both conditions. It would not have been surprising to see the priming effect change as a function of geographical distance from Liverpool. On the basis of how Hay & Drager (2010) explained their social class effect, it would have been expected that people who are less familiar with Scouse show less of a priming effect because they have fewer exemplars that can be activated by the prime. However, not only do we see an effect in the 'Manchester' group (where subjects should not have been biased to hearing Liverpool variants no matter where they live), but also this effect is actually in the direction which does not tie in with Hay & Drager's exposure-activation account. The interpretation of the possible relationship between frequency of remembrance and accuracy in perception given above, on the other hand, would predict the results this study has actually produced: Higher distance comes with

lower accuracy, only this time frequency of remembrance does not change due to lower salience, but to less frequent exposure to relevant variants. This account still does not explain why this effect is found for the *least* salient variable HAPPY and the *most* salient one /k/-lenition, but not the other two. At this point, I have no good explanation for this strange pairing which seems rather random. The issue of distance and familiarity clearly warrants further research, possibly based on a larger sample – both in terms of variables and participants – to arrive at a clearer and more detailed picture.

12.3 Non-social factors

12.3.1 Time held in memory

Let us now turn to independent variables which are non-social in nature. Hay, Nolan, et al. (2006) had originally worked with two different types of stimuli sentences because they wanted to investigate whether the priming effect depended on the time the subject had to hold the relevant sound in memory before they heard the synthesised answer tokens. However, they confounded position of the keyword with sentence length (sentences where the keyword appeared in the middle were always longer than sentences where the keyword was in final position), so it was not possible to tease these two factors apart. In the present study, care was taken to ensure stimulus sentences at least roughly contained the same amount of phonetic material (cf. §10.1.1). Results were ambiguous. Position of the carrier word within the stimulus sentence was found to be a significant predictor for both vocalic variables, HAPPY and NURSE, but not for the two consonants.

In the case of HAPPY, the reason for this is probably rather trivial and related to inconsistency in the synthesis of the answer tokens. As has been explained in §10.1.2, it was not possible to re-synthesise a satisfying vowel continuum out of the sentence-final HAPPY realisations. The continua synthesised for the corresponding sentence-medial stimuli were used instead. While the scale was the same and subjects always had a choice of options ranging from very central to very peripheral vowels, this had an unintended effect: For the sentence-medial stimuli (where HAPPY was naturally realised as [ɪ] by the speaker) the objectively best match was token 2, but in the sentence-final stimuli (where the speaker had [ə]) token 1 was objectively closest to what participants had heard. This means that subjects did not really behave differently with respect to sentence-medial and sentence-final stimuli. Rather, they always perceived HAPPY as slightly more fronted and raised than it actually was. In the case of sentence-medial stimuli

there was thus a slight preference for token 3 ([i]), whereas participants chose token 2 most often ([ɪ]), because it was already slightly more peripheral than the actually occurring [ə]. This explanation is very similar to how Hay & Drager (2010: 878–879; see also above) interpreted their age-related differences; the only difference is that in the present study the reference point people select the token against is not their own production, but the vowel occurring in the stimulus. Why subjects consistently perceived a vowel that was slightly more peripheral remains an open question, as it is very unlikely that participants' own production was even more central (across the board) than that of the Manchester speaker who recorded the stimuli (cf. Hay & Drager 2010: 878–879).

No such technical issues complicate the interpretation of the answers given for ɴᴜʀsᴇ stimuli. Participants produced a significantly lower average token number when the keyword containing the sound in question occurred in the middle of the sentence. The share of token 2 replies is also greater for these stimuli, so it can be said that subjects perceived sentence-medial stimuli slightly more accurately. This result is similar to what Hay, Nolan, et al. (2006) found. In their study, subjects had a tendency to answer with more New Zealand tokens when the keyword had been presented in the middle of the sentence. They are unsure about whether this is due to the fact that people have to hold the sound in memory (which might shift it towards their own production) or whether the shift is due to additional acoustic material that follows the keyword and provides further "phonetic cues which are associated with NZ" (Hay, Nolan, et al. 2006: 365). Priming did not have an impact on this effect. In the present study, there was also no significant interaction prime X position of keyword in the mixed-effects regression model, but if one looks at the raw data, the difference between priming conditions seems more pronounced for sentence-medial stimuli.

The author is tempted to accept Hay, Nolan, et al. (2006)'s explanation that following phonetic material (which was Mancunian in this study) shifts participants' perception towards the lower (Mancunian/standard) end of the scale. A shift towards subjects' own perception would also be possible, as it is very likely that the vast majority of perceivers had a central realisation of ɴᴜʀsᴇ. However, this account struggles to explain why the difference between sentence-medial and sentence-final stimuli seems to be driven almost exclusively by perceivers in the 'Liverpool' condition. It seems as if holding the sound in memory has the effect of making the stimulus sound even more Mancunian, but only to subjects who have been primed for Liverpool. What I think happens is that these participants were successfully primed to expect a Liverpool realisation of the ɴᴜʀsᴇ vowel. When they then get to the relevant vowel in the carrier word,

this realisation does not agree with the expectation that was created. Against the backdrop of their expectation, the vowel then sounds even more Mancunian compared to what it was *supposed* to sound like (cf. §12.5). This effect can be seen for sentence-final stimuli, too (subjects primed for Liverpool seemed ever so slightly less likely to perceive Scouse variants), but it grows tremendously in size for sentence-medial stimuli, presumably because subjects are confronted with further material that is in conflict with their expectation.

The amount of conflicting material then does not seem to be as important as the question of whether it occurs before or after the sound that needs to be memorised and then categorised (remember that both sentence-medial and sentence-final stimuli contained roughly the same amount of phonetic material). I am still reluctant to generalise from this result to claiming that larger priming effects are to be expected if participants have to hold the relevant sound in memory. First of all, this is not corroborated by previous research in sociophonetics (Hay, Nolan, et al. 2006; Hay & Drager 2010). Secondly, it would also contradict findings in social psychology, where studies have shown that subjects are most likely to apply 'automatic' processing (which relies on stereotypes) when they have to react very fast – 'controlled' processing, which is more conscious and objective, can only occur when subjects are given enough time (cf. Petersen & Six 2008: 33–34). When the keyword appears in the middle of the sentence, participants should therefore have to rely less on their stereotypes as they have more time to process the stimulus. In any case, there does not seem to be a straightforward explanation why the present study only found such an effect for NURSE, but not for lenition of /k/. It is clear that there is something interesting going on here, but it remains to be seen whether the effect can be replicated in different contexts.

12.3.2 Stimulus order

Another effect which only surfaced in the responses pertaining to NURSE sentences is the impact of stimulus order. In these data there was almost a statistical trend for subjects to answer with a lower number token the further they progressed in the experiment. Answers thus became more Manchester-like towards the end of the test. There is at least some evidence that the priming effect is largest at the beginning and diminishes as people work their way through the stimuli, but it has also been pointed out that there are a number of things which caution against making too much of this. Most importantly, priming only has a significantly different effect in the raw data, whereas the mixed-effects regression did not find an interaction of prime and stimulus order. Both Hay, Nolan, et al. (2006) and Hay & Drager (2010) found a similar effect and provided a number

of possible explanations: (1) The further they were into the experiment subjects "relied more on the most frequently activated exemplars", (2) they "relied more on exemplars representing their own speech", and (3) they were "getting increasingly used to the speaker's voice", arguing that the trend "may reflect an increase in accuracy" (Hay & Drager 2010: 881–882).

Especially the last point seems to be a reasonable interpretation of what might be going on: Participants just need a certain time to home in on token number 2. In an earlier study I have interpreted an impact of stimulus order in one (!) condition as evidence for a priming effect which is then corrected by acoustic material that is in conflict with the prime (Juskan 2011). This account would lead us to only expect such an effect for the 'Liverpool' condition, because if participants are primed for Manchester there is no 'wrong' expectation that could be corrected by diverging acoustic input. In the present study, however, there is an order effect in *both* conditions, even if it is slightly weaker in one of them (the fact that the effect is greater in the 'Manchester' group is not necessarily a problem; cf. §12.5). An explanation based on 'getting used to the speaker', on the other hand, works fine for both conditions. Just as with position of the keyword the question remains as to why there is only an effect for NURSE, but not for the other three variables (or at least for /k/ as the other salient one). It should also be borne in mind that the effect is not even quite a statistical trend, and might, in fact, be nothing but an artefact that could disappear in a larger sample.

12.3.3 Phonological environment

For the two consonants there was an additional predictor because stimuli could be distinguished with respect to whether the variable occurred as the last sound of the carrier word or whether it was presented in an intervocalic context. Phonological environment was a (highly) significant predictor for both consonantal variables. The effects were also nearly equally strong, but the direction was different: While for velar nasal plus subjects were more likely to choose higher number tokens when the variable was at the end of the carrier word, they actually showed a preference for lower number tokens in the same context for /k/ stimuli. For word-final stimuli, participants were thus *more* inclined to perceive Liverpool variants of /ŋ(g)/ and *less* willing to hear lenited realisations of /k/. Mixed effects regression did not find an interaction of prime and phonological environment for either variable, but graphical inspection of the raw data at least suggests the priming effect to be slightly stronger when the sound in question occurs word-finally.

As far as velar nasal plus is concerned, I can only offer a very tentative explanation of the results. Earlier in this book (cf. §11.3.3) I have presented the idea that this variable might actually be slightly more salient in _# environments because of the variants that are often encountered there. Knowles (1973) already found that /ŋ(g)/ is often realised as [ŋk] at the end of words in Liverpool (cf. §3.3.1) and, while I did not code for this, these variants are also commonly found in my own data. When /ŋ(g)/ occurs in-between vowels, devoicing does not usually take place because switching voicing off between two voiced sounds would be uneconomical. The [ŋk] or even [ŋkʰ] variants in word-final position could be somewhat more salient because they are phonetically even more different from the standard realisation [ŋ] than [ŋg] is. For intervocalic [ŋg] realisations the only cue for the plosive is often a (sometimes very subtle) burst since there is frequently no real silence phase because voicing is maintained all throughout. Also, in this context, a burst that is low in amplitude might be 'misinterpreted' as a somewhat too rapid release of the velar nasal and might therefore not attract as much attention.

In the case of /k/ lenition there is more evidence to base an interpretation on. Here, I think, it is a lot less controversial and speculative to claim that social salience is at least part of the explanation. This is because the perception data mirror quite nicely what was found in production. Liverpool speakers showed less style shifting for /k/-lenition in intervocalic contexts, presumably because these variants can be 'justified' phonetically and are therefore potentially slightly less salient in this environment (which is why lenited variants are especially frequent, cf. §7.2.2). Going from a vowel (with no real obstruction of the airstream) to the closure of a plosive (complete blockage of the airstream in the oral cavity) and back to another vowel (no blockage again) is the most extreme phonetic difference possible between two linguistic sounds. Lenition (to fricatives) reduces this contrast (and the articulatory effort) considerably, and replacing a plosive with a fricative is thus an instantiation of the principle of economy and a 'natural' thing to occur.

When /k/ appears in _# environments, on the other hand, participants show somewhat lower lenition rates (though still high in absolute terms) in production, which indicates that Liverpool variants are a bit *more* salient in these contexts. The fact that, in the same context, the priming effect in perception is also slightly greater is therefore highly interesting and relevant with respect to the main hypothesis of this study. Even though the statistical grounding is not too robust and conclusions will therefore have to be tentative, these results can be taken as evidence that the salience of a variable might not only be able to explain general

trends in exemplar priming experiments, but can, in fact, serve as a (maybe *the*) crucial predictor that is also capable of shedding light on diverging results in sub-groups of stimuli (phonological environment) or subjects (social class).

While salience is a good explanation for different sizes of the priming effect in the two phonological environments, it does not straightforwardly predict why the average token number in word-final stimuli would be lower, i.e. why subjects were more likely to perceive the ultra-Mancunian/standard token 1 (plosive with burst but no aspiration) in this context. We can, in fact, apply the same explanation if we assume that higher salience in this context results in a (stronger) priming effect in the unexpected direction (cf. §12.5). Put simply, participants expect a lenited variant (even more so than in the other context), and the actual acoustic input therefore sounds even more standard/Mancunian than it actually is, because it is judged against a very 'Liverpool-like' baseline. Subjects could then be said to hyper-correct their perception more for word-final stimuli because the distance between the activated exemplars and the actual acoustic input is greater in this context due to stronger activation by the prime.

In this particular case, however, there is also an alternative explanation which has been hinted at earlier: The higher proportion of token 1 answers in word-final contexts might be to do with expectations based on common allophonic distributions. Word-final plosives in English are often unreleased, i.e. not articulated with an audible burst or friction. These unreleased realisations are much rarer in intervocalic position. If participants expected to hear a non-released /k/ at the end of a word because this variant is often encountered in natural language in this position then this might have biased them towards choosing token number 1 (which did have a release burst, but no aspiration) as this was the answer option which best corresponded to their expectation. When the /k/ occurs in the onset of syllables (as it does in intervocalic environments) unreleased realisations are less common and so subjects choose token 1 less often. In this framework, participants would actually have been primed, but not with social information regarding the speaker, but rather by the phonological context the variable was presented in. I will call this allophonic priming, because of its grounding in the language system and its largely regular, i.e. rule-based realisations.

Both accounts predict the results equally well, so it is not really possible to decide which one is preferable on the basis of the current dataset. In both cases, however, we are looking at another piece of evidence for an exemplar account since perceivers seem to be influenced by prior experiences that are different depending on the context, and which must have been stored separately. However, it is an important insight in its own right that these different results might have

been brought about more directly by the phonological environment itself (via its typical allophonic /k/ realisations) instead of indirectly through social salience attached to it. This could in fact mean that a second, unintended level of (allophonic) priming was present in the research design. While it has just been outlined that, in this study, the two effects (social priming and allophonic priming) would produce a shift in the same direction, it is easily conceivable that this does not always have to be the case. In a scenario where these two factors are at odds, i.e. where the phonological environment biases participants towards other variants than the actual prime that is investigated, they might produce considerable noise in the dataset. As far as I am aware, the priming potential of allophonic distributions has not figured prominently in the sociophonetic literature up to this date, but it is certainly an interesting avenue for future research. In any case, researchers should take care to avoid any potential conflicts between allophonic and other sorts of priming to make sure they actually measure what they mean to measure and do not draw their conclusions on the basis of 'corrupted' datasets.

12.4 Issues and limitations

12.4.1 The problem of velar nasal plus

This analysis, like any piece of research, comes with a number of short-comings and limitations, a few of which will be addressed in the following paragraphs. The aspect of this study which will probably strike the reader most as potentially problematic is the use of velar nasal plus as a variable in the perception test. This is because the realisation of <ng> clusters as [ŋg] is not restricted to Scouse, but is actually found in an area that is roughly delimited by the cities of Liverpool in the west, Manchester in the east, and Birmingham in the south (cf. §3.3.1). In retrospect, it might have been preferable to contrast lenition of /k/ with lenition of /t/, because these two phenomena are, in fact, both restricted to Liverpool, at least in their extent and particular patterning. In the interviews conducted for this study, lenition of /t/ turned out to be much less salient to Scousers than lenited variants of /k/ (for instance, lenition of /t/ was never explicitly commented on), but this information was not available until the perception experiment was well under way. In the literature, lenition of these two variables (and others) is most often treated as one single phenomenon and the label 'highly salient' is usually likewise extended to all lenited plosive realisations. However, lenition of /t/ would have posed another serious problem, even if it had been clear from the start that its salience is lower than that of /k/ lenition. This is because /t/ not only

lenites to the affricate /ts/ or the fricative /s/, but also to [tθ], [θ], [h], and [∅] (cf. §3.3.2), so direct comparison with /k/ (and the creation of equivalent stimuli) would have been challenging to say the least. When the perception experiment was planned and designed, velar nasal plus seemed to be the only Scouse variable that was consonantal, less salient, at least roughly comparable to /k/ in phonetic terms, and relatively restricted in geographical spread.

It has been mentioned in §10.1.1 that the speaker from Manchester who recorded the stimuli also naturally produced [ŋg] realisations in the relevant carrier words. Priming would therefore not have been necessary to 'make' participants hear [ŋg] variants. Rather, it would have sufficed to let subjects perceive 'objectively' since these pronunciations would actually have been present in the stimuli. This was not the case because the stimuli for velar nasal plus had been edited appropriately. The speaker recorded the sentences using her native [ŋg] realisation, but the plosive was cut from the material before it was used in the perception test (cf. §10.1.3). The point of departure was therefore the same as for the other three variables: If subjects reported having heard the Liverpool variants 3 or 4, they must have been biased in their perception, because the realisation that actually occurred in the stimuli was in fact (close to) the standard or Mancunian one. This does not solve the issue of the priming categories used. Since velar nasal plus is present both in the speech of Liverpudlians and Mancunians, priming subjects for Liverpool or Manchester should, in fact, produce the same results: Participants should be biased towards hearing [ŋg] in both cases.

I would like to argue that this is not, in fact, the outcome that should be expected, due to the general salience of the Scouse accent as a whole in the linguistic landscape of the United Kingdom. In his 2007a PhD thesis Montgomery investigated laypersons' perceptions of several northern English varieties. One of the tasks he asked participants to complete was to draw in dialect areas on a (largely) blank map. His results were that, with a rate of almost 58%, Scouse was the most often recognised dialect region in his sample, whereas only about one participant in four drew and labelled a Manc area (cf. Montgomery 2007a: 194). Participants in his study also provided more "characteristics" (evaluations, comments, stereotypes) of Scousers than they did for Manc speakers. The difference was particularly pronounced with respect to linguistic features, where the Manchester area received only a single comment overall (cf. Montgomery 2007a: 246–252). Given that Manchester was not even recognised as a dialect area "on its own" in previous linguistic studies, Montgomery (cf. 2007a: 214–215) rightly points out that Manchester has gained in cultural salience in the last 2–3 decades (to a large extent probably due to pop music), but it seems quite clear that

Scouse as the "most salient" (Montgomery 2007a: 216) accent area is still much more present in people's minds.

The argument for including velar nasal plus in this study was thus that if Scouse as an accent is so much more culturally salient (and also stigmatised, as can be deduced from the explicit comments of people) than Manchester English, then priming for Liverpool should also have a stronger effect than priming for Manchester. In other words, it was hypothesised that participants either had no stored exemplars indexed with 'Manchester' at all (in which case priming them for Manchester would have had no effect at all), or, if they did, their number would be significantly smaller and/or activation of these exemplars would be weaker because the category is cognitively less present in general (in which case there would have been a priming effect for both 'Liverpool' and 'Manchester' in the same direction, but less strong in the latter case).

Remember that there almost *is* a priming effect in the velar nasal plus results (a statistical trend in the mixed-effects model), which is evidence that there is something going on between priming conditions. This constitutes post-hoc support for the reasoning presented above and shows that sticking to velar nasal plus as a variable in this experiment is justified. In fact, velar nasal plus shows more priming than was expected even when production data from Liverpool speakers (who showed some – slightly inconclusive – signs of salience) are taken into account. It seems as if velar nasal plus is more salient to outsiders than previously thought.

12.4.2 Comparability with previous research

As has been mentioned in several places already, this study owes a lot to previous research, especially Niedzielski 1999; Hay, Nolan, et al. 2006; Hay, Warren, et al. 2006; Hay & Drager 2010, whose results it was meant to replicate and contextualise. A number of characteristics of the present study, however, might reduce comparability with these papers.

For instance, most previous work has confronted subjects with resynthesised vowel tokens presented in isolation. This was deemed impractical for my own study, primarily because subjects were going to have to match consonants as well as vowels. In the case of /k/ in particular it would have been quite difficult to distinguish variants in perception as the tokens used in this experiment differ (not exclusively, but mostly) in the percentage of aspiration/friction in relation to the closure/silence phase. The duration of silence, in turn, can only be processed if there is acoustic material preceding said silence. Intriguingly, Hay & Drager (2010: 887–888) wonder about whether priming should even "be revealed in the

task that [they] have employed" because "perception of the [answer] continuum should be affected in the same way" as perception of the stimuli words/sentences. They furthermore speculate that priming might affect perception of the synthesised tokens less only because they are not "word-embedded, natural stimuli", and can therefore be "processed for what they are". In the present study, this does not seem to have been a problem, though, maybe because the resynthesised tokens still sounded artificial enough, even when they occurred within a word.

The second important difference in methodology is the sex of the speaker who provided the stimuli. Hay, Nolan, et al. (2006) and Hay & Drager (2010) had a male speaker record their stimulus sentences, but in the present study the speaker was female. This might not appear to be a major issue at first glance, but it could actually make a difference in the specific framework of this particular analysis. The two variables where the most pronounced priming effects were found (NURSE and /k/-lenition) are not only very salient but also highly stigmatised. It is very well possible that subjects either do not actively associate women very strongly with stigmatised variants, or that they have less stored exemplars that contain these variants *and* are indexed with 'female' (because these variants are, in fact, less frequent in the speech of female speakers in real life). It is therefore conceivable that stronger priming effects could have been found in this study if a male instead of a female speaker had recorded the stimuli. Further research replicating this study with a male speaker, or, perhaps preferably, an androgynous voice and a 2x2 priming scheme ('Liverpool female', 'Liverpool male', 'Manchester female', 'Manchester male') would be necessary to shed more light on this question.

12.4.3 The issue of frequency

The impact of frequency on the results of this study is rather unclear. An effect could only be found for happγ and /k/-lenition (which seems to be a strange pairing, given that the former is the least and the latter the most salient variable in this sample), and even in these cases it was not very robust, statistically speaking. What is more, the effect was actually in different directions: For happγ, percepts become *more* Liverpool-like, while, for /k/, stimuli they become *less* Liverpool-like with higher frequency of the keyword. Even weak frequency effects of this sort can be seen as general evidence for episodic accounts of language processing, but beyond that, I have no explanation to offer at this point that would account for these diametrically opposed results in a straightforward manner. It is likely that the potential impact of frequency is obscured in this analysis because the method employed was not primarily conceived to investigate frequency in the first place and is therefore only partially suited to do so. The issue will remain un-

til this (or a similar) study is replicated with a larger set of carrier words that constitutes a more detailed and more representative sample of the frequency range actually encountered in naturalistic language. However, this hypothetical study will most likely have to focus on one phonological variable only in order to keep the expense of time manageable for participants.

12.4.4 Size of the priming effect

As a last caveat it should be noted that although many of the priming effects reported on in this analysis were found to be statistically robust the importance of these effects should not be exaggerated. In almost all cases, there was a preference for the acoustically most accurate token number 2, or it was at least a close runner-up for first place. For the two salient variables, token 2 accounted for at least 60% of answers. What this means is that people were not fooled in the majority of cases. While their perception can be manipulated, this manipulation has limits. We *can* create a sort of penchant, but this does not mean that perception disregards the actual speech signal completely. I would like to stress that this also holds true of previous research: In Hay & Drager (2010)'s study, for instance, participants also show a clear preference for the acoustically closest or a very similar token, at least as far as the main variable of the study is concerned.

12.5 Direction of priming

The most important problem that the results of my perception experiment pose has so far been avoided, because it deserves a more detailed discussion and should therefore be treated separately. I am referring to the fact that whenever a statistically significant priming effect was found in the data, said effect was in the unexpected direction. Participants who had been led to believe the speaker they were going to listen to was from Liverpool were *less* likely to perceive variants that are typical of Scouse speech. This seems rather strange and calls for an explanation.

12.5.1 The problem of replicating priming experiments

First of all, I would like to point out that the present experiment is not the first sociolinguistic study that fails to closely replicate Niedzielski's, Hay, Nolan, et al.'s, and Hay & Drager's findings. In fact, at least two studies have looked at different variables in different locations and have not found a social priming effect at all.

Squires, for example, also looked at social priming in the context of language processing, but in contrast to most other studies in this area she did not focus on a phonological variable, but instead investigated the perception of subject-verb agreement. The two syntactic variables whose processing she analysed are NP+*don't* and *there's*+NP, both of which occur with singular and plural NPs (e.g. *the truck/trucks don't run* and *there's a truck/trucks in the driveway* (cf. Squires 2013: 206). For both features, usage of the non-standard variant can be linked to social class, but more so for invariant *don't*, the feature that is also more stigmatised in American English (cf. Squires 2013: 207–208). Participants were played recordings of ambiguous frames (___ *don't like it*; *there's* ___ *showing*) that occur with both singular and plural NPs in actual speech. The NP was replaced by white noise in the audio stimuli and the subjects were asked to indicate what they had 'heard' by selecting a visual representation of a singular or a plural NP (such as one bird vs. several birds). Participants were primed with the help of high- and low-status speaker photos that were shown while the audio stimulus was playing (cf. Squires 2013: 210–211). The hypothesis that high-status photos would favour standard responses, while low-status photos would decrease this rate was not borne out: "the social status of the target photo did not have an effect on sentence perception" (Squires 2013: 216).

It could well be that "social information simply does not affect morphological or syntactic perception in the same way that it does speech perception" as Squires (2013: 229) puts it. However, conflicting evidence also exists for phonological variables. Lawrence (2015), for instance, has looked at perceptions of BATH and STRUT. Both vowels are "widely acknowledged as highly salient markers of regional identity in British English", more specifically they divide England into a northern and a southern part: Southern speakers usually realise BATH as [ɑː] and STRUT as [ʌ], whereas speakers from northern England usually have [a] in BATH, and [ʊ] or [ə] in STRUT words (cf. Lawrence 2015: 1). Lawrence had his stimuli recorded by a speaker from Sheffield, resynthesised 6-step vowel continua and played these to listeners who were speakers of Southern Standard British English. Half of them were told the speaker they were listening to was from 'Sheffield, Northern England', while the other half were told the speaker was from 'London, Southern England' (cf. Lawrence 2015: 2–3). No significant priming effects could be found, neither for BATH, nor for STRUT. Lawrence (cf. 2015: 4) concludes that "the influence of social information on linguistic perception may be more limited than has been previously suggested".

Note that, at this point in time, it is not completely uncontroversial whether we *should* expect priming effects to be identical or at least similar in different

studies. Cesario (2014: 45), for example, claims that "the expectation of wide-spread invariance in priming effects is inappropriate". He argues that in order to replicate a study in the first place, we need to know which features "must be reproduced exactly for a replication attempt to be informative" and goes on to explain that we need to have "relevant theories that tell us that these features should matter" (Cesario 2014: 42). In his opinion, though, theories of priming are not yet advanced and sophisticated enough due to the "relatively young state of priming research". In consequence, researchers trying to replicate a study might unwillingly change "some critical feature of the experimental context" because it is – wrongly – "deemed irrelevant" (Cesario 2014: 43). This 'error' could, for instance, simply consist in "sampling from a population that differs markedly (…) from the population sampled by the original researcher" (Cesario 2014: 43). This might well be the case if samples in the different studies are from the US, New Zealand, and Britain.

12.5.2 Hay et al.'s explanation for inverted effects

In this study, however, the problem is not the lack of a priming effect, but the fact that it is reversed. Interestingly, both Hay, Nolan, et al. (2006) and Hay & Drager (2010) had to face the same issue in a sub-sample of their data. While a robust priming effect was found for the entire dataset in both cases, closer inspection revealed that not only was the trend "most strongly carried by the female participants" (Hay & Drager 2010: 875), but it was also actually *reversed* for the males (cf. Hay & Drager 2010: 876–877).

The authors hypothesise that this gender difference could be due to "differences in attitude". They argue that (a) there is a "fierce sporting rivalry between New Zealand and Australia", (b) that sport is the most important "cultural marker of nationalism" in New Zealand, and that (c) this is primarily a male domain. Hay & Drager deduce that male New Zealanders are "more likely to have negative associations with Australia, whereas females may have more positive (or neutral) associations". In consequence, women behave as expected when the concept 'Australia' is invoked and shift towards Australian exemplars (the authors liken this to accommodation in production). When men, on the other hand, are primed for 'Australia' they not only activate their Australian exemplars but also their negative associations with that country. In an attempt to disassociate with Australia (comparable to speech divergence), they then shift towards New Zealand exemplars (cf. Hay & Drager 2010: 884–885).

This is an interesting explanation, and one which at first glance appears to be neatly transferable to my own results. After all, it has been pointed out repeatedly

in this book that Liverpool English is one of the most heavily stereotyped varieties in Britain. Attitudes towards the city itself are also still widely negative, and dominated by concepts such as crime, poverty, deprivation, and decay. It makes perfect sense to assume that British listeners, when confronted with the category 'Liverpool', do not only activate any Scouse exemplars that they may have stored, but also their negative attitudes towards Liverpool itself, and that they may then wish to disassociate from the city and activate their non-Scouse exemplars even more strongly. All the same, this account has one crucial shortcoming: It only works for the outsiders. Whenever we find a priming effect in the responses provided by subjects from Liverpool itself, however, the shift is *also* in the unexpected direction, i.e. towards *less* Liverpudlian variants when the prime was 'Liverpool' (cf. the discussion at the beginning of Chapter 11). Liverpool speakers are clearly *not* disassociating from their city. Quite the opposite is true. Especially younger Scousers are rather happy to use Liverpool features to mark their local identity, as has been shown in the part of this book that is concerned with production data. Despite some linguistic insecurity, overt comments about Liverpool and its accent essentially tell the same story (cf. §8.3). Disassociation does therefore not seem to be a realistic option in explaining the reversed priming effect, at least not for Liverpool participants.

12.5.3 Assimilation and contrast effects

Research in social psychology provides an alternative explanation in the form of 'assimilation' and 'contrast' effects[1]. The priming effects we know from the literature on the integration of social information in language perception are instances of assimilation effects, where "[s]ubjects primed with exemplars of a particular category are more likely to use that category in evaluating a subsequently presented category-relevant stimulus" and to classify this stimulus "as an instance of that category" (Herr 1986: 1106–1107). Consider, for example, subjects primed with the concept 'Australia' who are then more likely to perceive Australian vowels. Also possible, however, are so-called contrast effects, where the outcome of priming can be described as "judgments inconsistent with, and opposite in nature to, the primed category" (Herr 1986: 1107). I am going to argue that the results of the present study (participants primed for Liverpool perceive *less* Liverpool-like tokens) can be understood as an example of a contrast effect.

In Herr's opinion, the crucial factor that determines whether priming will result in an assimilation or a contrast effect is the extent to which the primed cat-

---

[1]Heartfelt thanks go to Andrew MacFarlane for pointing me to the relevant studies.

egory and the stimulus overlap. When the prime is a "moderate category" and the stimulus is "ambiguous", then the stimulus "should in fact be judged as an instance of that category". If, on the other hand, the category used for priming is "extreme", then the "ambiguous target should not be categorized within the primed category" because there is little or even no match between the stimulus and the prime. The prime *will*, however, act as sort of a cognitive "anchor" for evaluating the stimulus (cf. Herr 1986: 1107). In other words: When the stimulus is reasonably similar to the primed category, then subjects will classify the stimulus as an instance of the *same* category, priming thus results in an assimilation effect. If, on the other hand, the perceived distance between the prime and the stimulus is too great because there is (next to) no overlap, perceivers will not only not categorise the stimulus in the same cognitive bin, but they will use the primed category as the 'standard' value and consequently shift the stimulus towards the other end of the scale because it is directly compared (or contrasted) with the prime.

Herr illustrates this principle quite impressively with an experiment manipulating subjects' expectations of 'hostility'. In this test Herr primed participants with the help of famous people that had, in pretests, been revealed as representing different levels of hostility. The primes fell into one of four categories: "extremely nonhostile" (e.g. *Pope John Paul* or *Santa Claus*), "moderately nonhostile" (*Robin Hood*, *Henry Kissinger*), "moderately hostile" (*Alice Cooper*, *Menachem Begin*), and "extremely hostile" (*Dracula*, *Adolf Hitler*). Each participant was confronted with one of the lists just mentioned (which consisted of four names each) in the form of a matrix of letters puzzle where the names from the list had to be identified. After priming, subjects were given a description of a fictitious person ("Donald") to rate, whose behaviour was ambiguous and could be classified as either hostile or non-hostile (cf. Herr 1986: 1108). Results of this experiment were pretty clear and as expected: There was an interaction of prime and its "extremeness". Participants primed with moderately non-hostile *or* extremely hostile exemplars rated "Donald" as less hostile, and more friendly and kind than did subjects who had been exposed to the moderately hostile *or* extremely non-hostile category (cf. Herr 1986: 1109). In other words, subjects rated the fictitious person as similar if the prime was moderate, but reversed the effect when priming used "extreme" categories; compared to, say, Adolf Hitler, Donald actually seems to be a pretty nice person. Obviously, this explanation only works in an episodic framework, as it necessarily requires the presence of stored exemplars that can act as reference points.

I believe the very same process is behind the results of my own perception test. The only assumption that needs to be accepted to explain the direction of priming as a contrast effect is that the prime used in this study was what Herr calls "extreme", and this does not seem too far-fetched. Actually, it is not at all implausible to think that trying to make listeners perceive a Manchester voice as a Liverpool one is pushing the whole affair too far. What seems to have happened is that the Manc variants of the two salient variables present in the recordings are phonetically too different for British listeners to categorise them in the same category as the Scouse variants, even if these exemplars *have* been activated. Priming does still have an effect, though: Perceivers use the primed category 'Liverpool' as the baseline that the acoustic input is categorised against, so evaluations are shifted towards the other end of the scale. The result is the contrast effect that was identified time and again in this study: The 'Liverpool' speaker sounds even more Mancunian than she would anyway.

The same reasoning might also explain why subjects chose the slightly-Scouse token 3 much more often for NURSE stimuli than for /k/ ones. It is true that, at least for English, vowels "on the whole carry more responsibility than consonants in determining differences between accents" (Foulkes & Docherty 1999: 12), which would make them more liable to carry social meaning, all other things being equal. However, they also form a natural continuum without clearly delimited borders (cf. Foulkes & Docherty 1999: 12), which could make it generally harder to unambiguously categorise a specific token as an instance of category A or B. Subjects might therefore be more tempted to at least occasionally categorise the perceived variants as Scouse, because vowel categories have more fuzzy boundaries per se, compared to the realisations of /k/ which are thought of as more categorical in articulatory terms (plosive, affricate, fricative) to start with. Herr's framework could also provide further explanation for the class X priming interaction in the results for /k/ lenition. Middle class subjects showed a contrast effect. Working class participants exhibited a trend in the other (i.e. expected) direction. This is rather speculative because there were very few observations for working-class perceivers, but if this trend solidifies and becomes a significant assimilation effect in a larger dataset it could well be because working-class subjects are less aware of and sensitive to social differences in language use, so the prime 'Liverpool' might actually be less extreme for them than it is for middle-class perceivers.

Finally, assimilation and contrast effects can also offer an alternative explanation for Hay, Nolan, et al. (2006)'s and Hay & Drager (2010)'s gender differences. The authors argue that women behaved as expected because their associations

with Australia are positive or neutral, whereas men have more negative attitudes due to the sporting rivalry and wish to disassociate from the invoked concept 'Australia'. It is equally possible, however, that the gender differences that were found in these studies are ultimately caused by different levels of awareness of variation in the test variable, rather than different attitudes. Normally, we would expect women to be more aware of a socially salient variable, but in this case it might actually be the men because they are more invested in the national rivalry based on sport. If the difference between New Zealand English and Australian English is only moderately salient for women, they should show an assimilation effect (which they did). For men, on the other hand, the distinction is very (!) salient. If their categories are (felt to be) more distinct or distant from one another, then the prime 'Australia' would actually be much more extreme for them than for the women. Possibly just as extreme as the prime 'Liverpool' was for the British subjects in my study, which might explain why both my participants and the male perceivers in New Zealand show a very similar contrast effect.

12.6 Summary and implications

So, what do the results of this study mean for exemplar priming in sociolinguistics and how do they relate to previous research?

First of all, priming does indeed seem to be limited to (highly) salient variables. Less or non-salient variables do not generate a priming effect or, at most, a very weak one. Secondly, priming might, on the whole and all other things being equal, turn out to work better with vowels than with consonants, at least in cases where the consonantal variants cannot be placed on a vowel-like continuum without comparatively straightforward boundaries (as in, for example, the [s]-[ʃ] continuum, cf. Strand 1999). This would also be in line with research from social psychology, which has found that we are most likely to look for the 'help' of stereotypes when making a certain decision (i.e. categorisation) is difficult (cf. Petersen & Six 2008: 28). Thirdly, the prime and the acoustic material to be categorised must not be too different, at least not if the 'goal' is to generate an assimilation effect.

The existing research by Niedzielski and Hay and colleagues, might in fact, only have succeeded in finding a priming effect because their studies (possibly unknowingly) fulfilled all three criteria. Both US English and Canadian English, on the one hand, and New Zealand English and Australian English, on the other, are accents that are – compared to the range of variation present in the anglophone world – relatively similar to one another (cf. Halford 2002: 31 for

Canadian English, and Hay, Nolan, et al. 2006: 354 for New Zealand). What is more, the variables used for testing (Canadian Raising for Niedzielski 1999, raising/centralisation of [ɪ] for Hay, Nolan, et al. 2006; Hay & Drager 2010) are very salient to speakers of these varieties, possibly because it is one of the few features (or even *the* feature) that distinguishes these varieties. And finally, they both involve rather fine-grained phonetic differences between variants, so perceivers might be more susceptible to the influence of priming because the task of categorising these stimuli is a comparatively difficult one to start with.

In Britain, for instance, the situation is very different because accents differ much more drastically from each other. As a consequence, priming can easily fall into the trap of 'overdoing it' by trying to suggest something to perceivers which is just too incredible to swallow, given that the actual phonetic material is too different from what it is supposed to be. This can then result either in a contrast effect, such as in the present study, or, possibly, in subjects' ignoring the prime altogether (like in Lawrence 2015) when they (sub-consciously) realise that it is not 'helping'. The violation of principle three above could also be behind the fact that social factors play a less prominent role in this study than was previously expected (gender does not turn up as a significant predictor, social class is only relevant for /k/). Maybe the conflict between the prime and the stimuli was so drastic that a lot of the potential impact of social factors was 'swamped' by the overwhelming contrast effect.

Finally, future research in this area should take care to avoid the pitfalls of adding noise or even a second and unintended priming effect to their data by not controlling for factors such as phonological environment and possibly also frequency of carrier words, although the evidence is less clear in the latter respect.

None of this is meant to imply that social priming in language perception is not real. It is now very well established that humans *do* store social information in long-term memory and later integrate it with acoustic data when they process linguistic input, and the present study has added further to the pile of evidence supporting this idea. After all, it did find a priming effect, even if it was not in the expected direction. What the results of this study might also be able to do is to put the whole priming paradigm into perspective: Priming works, but only in certain, very special contexts. While exemplar priming remains extremely interesting from a *theoretical* point of view, it is possibly a lot less important in *practical* terms than has previously been suggested.

Draft of 9th October 2018, 10:13

Sound change, priming, salience

This volume investigates the realisation and perception of four phonological variables in Liverpool English (Scouse), with a special focus on their sociolinguistic salience. Younger speakers' speech is found to be more local, but only for the two salient variables in the sample (NURSE-SQUARE and /k/ lenition), which appear to carry considerable amounts of covert prestige. Local variants of non-salient happy-tensing and velar nasal plus, on the other hand, are actually found to be receding, so at least to a certain extent Scouse also seems to be participating in regional dialect levelling.

The importance of salience is also obvious in the perception data, with only the two highly salient stereotypes generating robust effects in a social priming experiment (albeit in the unexpected direction). These results indicate that the investigated variables differ measurably not only in their use in production, but also in terms of how central they are to mental sociolinguistic representations of Scouse. They also tell us more about the way we process, store, and (re-)use sociolinguistic variation in perception. By defining likely contexts for significant priming effects they might finally even help in coming up with a more elaborate 'theory of priming' in the realm of sociophonetics.