Chapter 8: Discussion

8.1. Introduction

In this chapter we give a more detailed discussion of the results that were presented in chapter 7. The hypotheses proposed in chapter 4 were already supported by a rather extensive discussion, in which we paid attention to the predictions made by rule-based models on the one hand and exemplar-based models on the other. Thus, some of the discussions in this chapter recapitulate points made in chapter 4. In addition, we will try to account for the results which were not anticipated. If relevant, we will try to shed light on the question whether the results can best be accounted for by assuming a model of rule-based or of exemplar-based learning.

In our discussion of the results, we will more or less follow the same order as in chapter 7. Section 8.2 discusses the results on the effects of feature-related factors on the degree to which dialect features are acquired. Section 8.3 deals with the effects of speaker-related factors on the degree of success in (second) dialect acquisition. In section 8.4 we take a closer look at the results concerning the degree of overgeneralization of dialect features. Finally, section 8.5 presents conclusions.

8.2. Discussing the effects of feature-related factors on the degree of dialect knowledge/acquisition

8.2.1. Accounting for the effects of incidence

In section 7.2.2.1, we demonstrated that the incidence or type frequency of dialect features has a highly significant, positive effect on the degree to which the relevant features are realized correctly by the second dialect learners and the native dialect speakers (cf. tables 7.2 and 7.3). This implies that a feature is known/acquired more successfully if its incidence is high(er) (i.e. if it applies to a large number of words). This outcome confirms the hypothesis that was presented in section 4.2.3, i.e. that high(er) incidence leads to better acquisition.[1] In section 4.2.3, we pointed out that both rule-based and exemplar-based theory make the same prediction. Below, we recapitulate how both models can account for the positive effect of incidence.

Although frequency effects have not been the primary focus of rule-based theories, it is not impossible to account for the positive effect of type frequency on the degree to which dialect features are acquired. Within a rule-based framework, it is assumed that language learners form a (correspondence) rule to generalize over a number of forms that display, for

[1] Since there was a significant interaction of incidence with home language, we can draw conclusions that are specific for second dialect acquisition (as opposed to first dialect acquisition).

example, the same dialect feature (e.g. all forms that display deletion of an underlying /l/ and lengthening of the preceding vowel). From the perspective of second dialect acquisition, it can be argued that such a (correspondence) rule is a very economical way of learning: once the child knows the rule, it can apply this rule to all words that meet its requirements. Under this assumption, the positive effect of type frequency on the acquisition of certain dialect features can be accounted for as follows. It seems reasonable that it is easier for a child to form a (correspondence) rule when that rule applies to a large part of the lexicon: children are more often exposed to a dialect feature when that feature occurs in a (relatively) large number of words. As a result of this high incidence, it is easier for a child to derive a (correspondence) rule that generalizes over these different words. This leads to better acquisition.

As we argued in section 4.2.3, an exemplar-based model can also account for the positive effect of incidence on the degree of acquisition. The account runs as follows. Exemplars (i.e. memorized words and word forms) are grouped into associative networks (i.e. neighbourhoods or exemplar clouds) on the basis of similarity, and there are mutual associations among the exemplars. As a result of these associations, the activation of one item in the neighbourhood leads to the activation of all associated items. Hence, the larger the exemplar cloud, the more items are activated. Since these items all display the same patterns, the level of activation of the entire pattern (or feature) increases. As a result, the feature gains lexical strength, which, according to Bybee (2001:120), leads to a high degree of 'analysability'. Since features with a high type frequency or incidence involve large exemplar clouds, we may conclude that under an exemplar-based account, high incidence should have a positive effect on the degree of acquisition of a feature.

Summarizing, we may say that both rule-based theory and exemplar-based theory can account for our results with respect to the factor incidence.

Returning to our results, we also found (see section 7.2.2.1) that the positive effect of incidence on the degree to which features are realized correctly is stronger in the case of the control group than in the case of the second dialect learners. This result seems to indicate that incidence is important to the success in language acquisition in general. Apparently, the factors that play a role in the acquisition of the Maldegem dialect as a second language are the same as those that play a role in the acquisition of this dialect as a first language.

8.2.2. Accounting for the effects of competing variants

In section 4.2.2, we hypothesized that dialect features involving a small number of competing dialect/Standard Dutch variants (i.e. approaching a one-to-one relationship between SD x and dialect y), are acquired better than features involving a large number of competing variants. The effect of the number of competing dialect variants was tested in section 7.2.2.2.1, and the effect of the number of Standard Dutch variants was tested in section 7.2.2.2.2. In this section, we discuss the results.

Consider the situations represented in (1).

(1) L1/SD L2/DIA

(a) x → a, b, c

(b) a, b, c → x

In a situation such as (1a), the dialect learning child must map one L1 form onto three different L2 forms, whereas in situation (1b) it must map three different L1 forms onto a single L2 form. Our results in section 7.2.2.2.1 revealed that there was no significant main effect of the number of competing dialect variants (cf. 1a above) on the degree of dialect knowledge, and neither was there a significant interaction of this factor with home language (see table 7.4). On the other hand, the number of competing Standard Dutch variants (cf. 1b above) had a significant negative effect on the degree of dialect knowledge and interacted significantly with home language (see table 7.5; section 7.2.2.2.2). Hence, we can conclude that the situation represented in (1b) causes more learning difficulties than the situation in (1a). This result contradicts the findings from the literature that the acquisition of new phonological contrasts (cf. situation 1a) is more difficult than the suppression of 'old' (i.e. L1) contrasts that do not exist in the L2 (cf. situation 1b) (cf. Trudgill 1986; Archibald 1998).

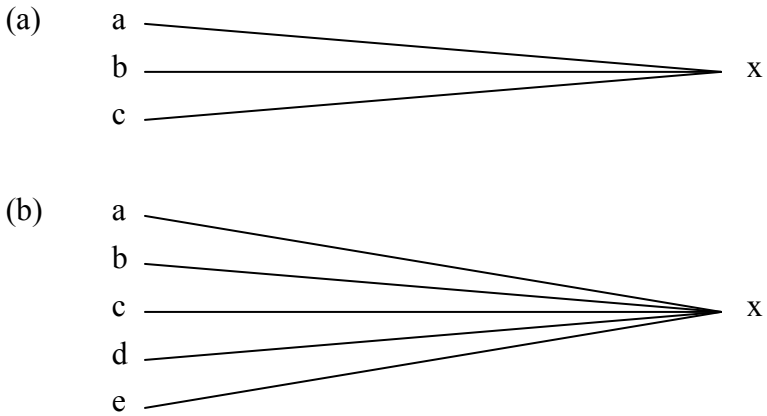
Now consider the situations represented in (2) and (3) below.

(2) L1/SD L2/DIA

(a) x → a, b, c

(b) x → a, b, c, d, e

(3) L1/SD L2/DIA



In section 7.2.2.2.1 we tested whether an increasing number of dialect variants interferes with acquisition (cf. 2a and 2b above), and in section 7.2.2.2.2 we tested whether an increasing number of Standard Dutch variants makes acquisition more difficult (cf. 3a and 3b above). In the following two sections, we try to account for our results.

8.2.2.1. Accounting for the effects of the number of competing dialect variants

In section 4.2.2, we hypothesized that features involving a relatively large number of competing dialect variants would be more difficult to acquire than features involving a relatively small number of competing dialect variants. However, on the basis of the results discussed in section 7.2.2.2.1, we cannot decide that this hypothesis is confirmed, since there was no significant effect of the number of dialect variants on the degree of dialect knowledge/acquisition.

8.2.2.2. Accounting for the effects of the number of competing Standard Dutch variants

In section 7.2.2.2.2, we discussed the effect of the number of competing Standard Dutch variants on the degree to which dialect features are acquired. The hypothesis with respect to this variable was confirmed by the results: there is a highly significant negative effect of the number of Standard Dutch variants on acquisition, which indicates that dialect features involving a relatively small number of SD variants are acquired better than features involving a relatively large number of SD variants (i.e., in situation 3b the dialect variant is acquired less successfully than in situation 3a).

The negative effect of this variable can be accounted for both from a rule-based and an exemplar-based perspective. The rule-based account runs as follows. In both the situations represented in (3), a second dialect learner knows precisely which words receive SD variant a, which ones receive SD variant b, and so on, because Standard Dutch is the learner's mother

tongue (L1). In order to acquire the dialect features involved (e.g. SD a ~ DIA x), the dialect learner has to learn the different correspondences involved (i.e. SD a ~ DIA x; SD b ~ DIA x; SD c ~ DIA x; etc.). When there are more competing Standard Dutch variants, more correspondences have to be formed, which may complicate acquisition. This explains why the number of competing Standard Dutch variants has a negative effect on the degree to which dialect features are known/acquired. So, dialect features will be acquired less successfully in situation (3b) than in situation (3a).

An exemplar-based account of these effects is also possible, and actually does not deviate very much from a rule-based account.[2] An important notion in the exemplar-based account is the degree of heterogeneity (or homogeneity) of exemplar clouds (see also section 4.2.2). As mentioned before, an exemplar cloud is heterogeneous when it contains items with sound x which can either refer to items with sound y or with sound z. In such a case, the items with sound y and those with sound z are competing forms (i.e. y and z are competing variants). The basic exemplar-based prediction is that acquisition becomes more difficult if there is a higher degree of heterogeneity of a cloud (i.e. more competing forms/variants). This prediction is independent of the question whether the exemplar cloud contains words or word forms with a Standard Dutch (L1) vowel or with a dialect (L2) vowel.

Suppose that there is a cloud consisting of words with the dialect vowel /ø/. This cloud is very heterogeneous, since it contains words with DIA /ø/ corresponding to SD /ɛi/ (e.g. *lijm* 'glue', *ijs* 'ice', *bijten* 'to bite'), to SD /œy/ (e.g. *duim* 'thumb', *huis* 'house', *buiten* 'outside'), to SD /o:/ (e.g. *zoon* 'son', *boter* 'butter', *door* 'through'), to SD /ø/ (e.g. *deur* 'door', *reus* 'giant'), and (a small number of) words with DIA /ø/ corresponding to SD /y/ (e.g. *duwen* 'to push', *schuw* 'afraid'). Hence, there are five competing SD variants (see section 6.4.6). Acquisition is predicted to be difficult in this situation. For example, the exemplar [bøtn̩] may refer to the word *bijten* 'to bite' (with SD /ɛi/), but can also refer to the word *buiten* 'outside' (with SD /œy/). This complicates acquisition.

From table 7.5 and figure 7.3, it became clear that the negative effect of number of Standard Dutch variants is much stronger in the case of the native dialect speakers than in the case of the second dialect learners (there was a significant interaction effect of n SD variants with home language). The negative effect on the acquisition of the second dialect learners was so weak that it was not visible in the graphical representation. The control group displays a similar, though much stronger, reaction to the influence of the factor number of Standard Dutch variants as the second dialect learners. The fact that native dialect speakers are strongly influenced by the number of Standard Dutch variants that correspond to particular dialect elements can be explained by the fact that all children involved in this study (including the

[2] The main difference between the rule-based and the exemplar-based account is that the former assumes the existence of intersystemic correspondence rules, whereas the latter assumes the existence of exemplar clouds.

dialect speakers) have a good proficiency in Standard Dutch, since it is the language variety spoken at school (and in other official situations). We cannot explain, however, why the dialect of the second dialect learners is much less affected by the number of competing SD variants.

8.2.3. Accounting for the effects of conditioning environment

We have argued that the degree of predictability of a dialect feature increases when the relevant feature is restricted to a conditioning phonological environment. Our hypothesis in section 4.2.4 stated that dialect features that are restricted to a conditioning environment will be acquired better than features without environmental restrictions. In section 7.2.2.3, we demonstrated that the results on the effects of conditioning environment support our hypothesis: there was a significant main effect, but no significant interaction effect with home language, so both the second dialect learners and the control group have a better knowledge of dialect features with a conditioning (phonological) environment than features without it (cf. table 7.6, figure 7.4).

We have argued that both rule-based and exemplar-based models make predictions that are consistent with our hypothesis. Within a rule-based framework (e.g. in the models of Auer 1993 and Taeldeman 1993), we can argue that dialect learners form correspondences between equivalent elements of their L1 and the L2. For example, learners may form a correspondence that states that SD /ɛi/ corresponds to dialect /a̟i/ (e.g. in *eigen* 'own', *kei* 'boulder'). However, this correspondence is of little use in facilitating dialect acquisition, since it is unpredictable to which words the correspondence applies: the correspondence SD /ɛi/ ~ DIA /a̟i/ only applies to a limited number of words, and is completely lexically determined. This means that even if the learner forms the correspondence, he still has to learn whether the correspondence applies word by word.

On the other hand, the correspondence between SD /ɛi/ and dialect /ɐ/ (e.g. in *kijken* 'to look', *zwijgen* 'to be silent') is restricted to positions before a velar or laryngeal consonant. From a rule-based perspective, we can argue that when there is a conditioning environment, a learner integrates this conditioning environment in the correspondence. The integration of this conditioning environment in the correspondence makes the correspondence more specific, which may lead to a better acquisition of the relevant dialect feature (and, as we have argued in section 4.4.4.3, to a lower degree of overgeneralization of the feature).

Recall from section 4.2.4, that exemplar-based theory basically predicts a positive effect of conditioning environment on the degree of acquisition as well. We argued that the concept of conditioning environment (as implemented in this study) can be related to the exemplar-based notion of phonological neighbourhoods, i.e. groupings of words (or morphemes or smaller units) in the mental lexicon which display a high degree of overlap in their

phonological make-up (e.g. the words *lijk* 'corpse', *rijk* 'rich', *dijk* 'dike' and *wijk* 'district' together constitute a possible phonological neighbourhood). When a child acquires the correct dialect variant of a word which belongs to a particular phonological neighbourhood, this dialect variant also spreads to the other exemplars in the neighbourhood. Since we have suggested that most words which group together in a phonological neighbourhood have the relevant phonological environment as a *conditioning* environment (although some words in the neighbourhood will have the environment only as one of many *possible* environments), we may conclude that conditioning environment has a positive effect on the acquisition of the correct dialect variant.

The positive effect of conditioning environment on the degree to which dialect features are known can therefore be accounted for by the idea of phonological neighbourhoods, provided that the phonological neighbourhoods involved (e.g. a neighbourhood organized around the biphone /ɛik/) largely coincide with the conditioning environments as we have defined them for each feature (e.g. "before velar or laryngeal consonant").

In section 4.2.4, we suggested that the effect of conditioning environment should preferably be considered in interaction with other factors, such as (average) token frequency, incidence, productivity, number of competing variants, etc. Below, we focus on the factors which may complicate the account of phonological neighbourhood effects given above.

Stating that a conditioning environment has a positive effect on dialect acquisition may be too simple a representation, especially from an exemplar-based perspective. We have argued that phonological neighbourhood effects most likely depend on the degree of heterogeneity of the relevant phonological neighbourhood, i.e. the degree to which all exemplars in the neighbourhood refer to the same dialect variant. We suggested that it would therefore be appropriate to calculate the interaction effect of conditioning environment with number of competing variants, under the assumption that the relevant phonological neighbourhood is determined by the same phonological conditions as the ones on the basis of which we determined the conditioning environment of the features involved in this study.

For example, in our research design, the feature SD /ɛi/ ~ DIA /e/ has the conditioning environment 'before velar/laryngeal consonant'. This conditioning environment covers words with /ɛik/ (e.g. *kijken* 'to look'*, strijken* 'to iron'*, slijk* 'mud'), as well as with /ɛiɣ/ (e.g. *krijgen* 'to receive'*, zwijgen* 'to be silent'). Hence, we could argue that there is a phonological neighbourhood with words containing either /ɛik/ or /ɛiɣ/. However, the notion of phonological neighbourhoods does not necessarily coincide with our notion of conditioning environment. For example, we could also consider the words with /ɛik/ as belonging to one phonological neighbourhood, and the words with /ɛiɣ/ as belonging to another neighbourhood. In that case, the number of possible competing variants (i.e. the degree of heterogeneity of a neighbourhood) changes. The words in the /ɛik/

neighbourhood would only refer to two dialect variants, i.e. to dialect /ɐ/ (e.g. *kijken* 'to look') and to dialect /iə/ (e.g. *eikel* 'acorn'), whereas the words in the /ɛiɣ/ neighbourhood would refer to four (competing) dialect variants, i.e. to dialect /ɐ/ (e.g. *zwijgen* 'to be silent'), to dialect /iə/ (e.g. *dreigen* 'to threaten'), to dialect /æi/ (e.g. *eigen* 'own'), and to dialect /i/ (e.g. *tijger* 'tiger'). So, the latter neighbourhood is more heterogeneous than the former.

The research design of this study did not start out with an exemplar-based model. Hence, we did not implement the factors of conditioning environment and number of competing variants in a way that is totally consistent with the exemplar-based assumptions. For example, /ɛiɣ/ and /ɛik/ are considered as constituting one conditioning environment. In addition, we implemented the number of competing variants independently of the factor of conditioning environment.[3] Because of these reasons, we should be cautious in drawing firm conclusions about the exemplar-based account of the effect of conditioning environment on the degree of dialect proficiency.

Nevertheless, we calculated the interaction effect of conditioning environment with number of competing variants on the degree of dialect acquisition (see section 7.2.2.7). This interaction will be discussed in section 8.2.7.1 (table 8.1).

Another factor of which we assumed that it might complicate the exemplar-based predictions about the effect of conditioning environment, was token frequency (see section 4.2.4). We argued that the effect of conditioning environment would be stronger when the relevant phonological neighbourhood mainly contained frequently used words. Therefore, we suggested that the interaction between conditioning environment and average token frequency per feature should be calculated. This interaction will be discussed in section 8.2.7.2.

Apart from the degree of heterogeneity of exemplar clouds (related to the number of competing variants) and the average token frequency per feature, there is another factor which we argued could complicate the effect of conditioning environment (according to exemplar-based theory), i.e. the distance between exemplar clouds. In section 4.2.4, we suggested that it is easier to acquire the correct dialect variant of words belonging to phonological neighbourhoods that are relatively remote in exemplar space, than in the case of neighbourhoods which are closely surrounded by other neighbourhoods. Since we did not base the design of our study on an exemplar-based model, we did not implement a factor which takes into account the degree of remoteness of phonological neighbourhoods. So, we cannot test all of the predictions made by exemplar-based theory.

Summarizing, we may say that the positive effect of conditioning environment on the degree to which dialect features are known, can be accounted for from a rule-based as well as

[3] In this way, we obtained six competing variants for the features involved in the /ɛi/-paradigm (see section 6.4.5).

an exemplar-based perspective.[4] However, interactions of conditioning environment with other factors must be taken into account, particularly when testing the predictions made by exemplar-based models (see section 8.2.7).

8.2.4. Accounting for the effects of token frequency

In this section, we first account for the effect of token frequency on the degree to which words are realized correctly (section 8.2.4.1). Next, we account for the effect of average token frequency on the degree to which dialect features are known or acquired (section 8.2.4.2).

8.2.4.1. Token frequency at the level of the word

In section 7.2.2.4.1, we demonstrated that the words from our word list with a high frequency of usage (i.e. a high token frequency) were more often realized with the correct Maldegem dialect variant than the words with a low(er) token frequency (see table 7.7 and figure 7.5). This result confirmed the hypothesis that was proposed in chapter 4, as far as the degree of dialect knowledge is concerned; since there is no significant interaction between token frequency and home language, we cannot draw any conclusions about the effect of this variable that are specific for second dialect acquisition in particular.

Below, we discuss whether and how these effects of token frequency can be accounted for within rule-based and exemplar-based frameworks (see also the discussion in section 4.2.5).

Effects of token frequency have been noted in processes of language change. Consider the well-known example of Kloeke's *huis-muis*-map (Kloeke 1927). Kloeke shows that the diphthongization of Wgm. *û* into modern Standard Dutch [œy] (via [y]) proceeded gradually through the lexicon (see also Janssen 1941). He observes that this sound change affected words that were frequently used in interurban communication (e.g. *huis* 'house') earlier than infrequently used words (e.g. *muis* 'mouse') (see also Taeldeman 2006b). Traditional, rule-based accounts of processes of language change recognized the existence of frequency effects, but they were not really concerned with an explanation of these effects. Actually, rule-based models of language change cannot easily account for the fact that a sound change does not usually affect all words at the same time.[5] Likewise, effects of token

[4] Since there was no significant interaction of conditioning environment with home language, we cannot draw any conclusions about the variable 'conditioning environment' which are specific for second dialect acquistion alone.

[5] This does not mean that rule-based models cannot account for lexical diffusion effects at all. Kiparsky (1995) shows how the lexical diffusion of *ū*-shortening (e.g. in *cook*, *shook*, etc.) can be explained as the extension of a structure-building rule: the rule's context (i.e. [-anterior] ____ [-anterior, -coronal]) is extended so that words with the environment ____ [-anterior, -coronal] (e.g. *took*, *book*, etc.) or with the environment [-anterior] ____ (e.g. *good*, *could*, etc.) also fall under its scope. However, this process takes place slowly and word by word, since some words can be prespecified with two moras in the lexicon so that they escape the shortening rule (at

frequency are rather problematic for rule-based models of language acquisition. The basic assumption of rule-based models is that the language learner generalizes over a number of L2 forms by forming rules. When applied to our study, this means that the second dialect learner forms correspondences between his L1 and the L2. Once the dialect learner has acquired such a correspondence rule and its conditions (e.g. when he has been confronted with 'enough' words that show a given dialect feature), he should be able to apply this correspondence to all words that meet the conditions, no matter how frequently these words occur. It is therefore not easy to explain why a dialect learner sometimes applies a particular correspondence rule in high-frequency words, whereas he does not do so (yet) in low-frequency words.

Recall from chapter 3 (section 3.3.3) that much attention has been paid to frequency effects within exemplar-based theories. Our results on the effects of token frequency can therefore best be accounted for within an exemplar-based framework. The idea of an exemplar-based model like that of Bybee (2001) is that words that are used in the input are directly mapped onto the most similar memory representations in the mental lexicon (i.e. the nearest neighbours). As a particular word is more frequently used in the input, the mental representation on which this word is mapped grows stronger and thus becomes easier to access. However, infrequently used words may fade away and are therefore difficult to access. This can account for the fact that high-frequency words are more often realized with the correct Maldegem dialect variant than low-frequency words (see figure 7.5).

Consider the Maldegem dialect feature 'SD /ɛi/ ~ DIA /æi/': learners are more often exposed to a frequent word such as *meisje* 'girl' (token frequency = 310, see appendix 3), also in its dialect form, than an infrequent word such as *reiger* 'heron' (token frequency = 1). Following Bybee, we can argue that the lexical representation of *meisje*, which is stored with idiosyncratic information (e.g. information about the dialect pronunciation, viz. [mæiskə]), is frequently activated, whereas that of *reiger* is only sporadically activated. The dialect form of *meisje* is therefore more easily accessed than that of *reiger*. This explains why the dialect variant (i.e. /æi/) is used more often in *meisje* than in *reiger*. A purely rule-based model predicts that a dialect learning child forms a correspondence between Standard Dutch /ɛi/ and dialect /æi/ and is able to apply this correspondence to all words it is exposed to, without making a distinction between high- and low-frequency words. This, however, is not the desired result.

To conclude, the fact that token frequency plays a role in the degree to which the dialect features are realized correctly, is an argument in favour of an exemplar-based model of language learning.

---

first). Lexical diffusion occurs because these words are slowly "'regularized' through loss of the prespecified length in their underlying representations" (Kiparsky 1995:648).

8.2.4.2. Average token frequency per feature

In section 7.2.2.4.2, we showed that average token frequency has a positive effect on the degree of dialect proficiency/acquisition. This positive effect implies that dialect features which mainly occur in frequently used words (i.e. which have a high average token frequency) are known or acquired better than features which mainly occur in infrequently used words (see table 7.8 and figure 7.6).

The exemplar-based account discussed in the previous section concerned the effect of token frequency on the degree to which individual words are realized with the 'correct' dialect variant. Exemplar-based models are also very suitable to account for the positive effect of the average token frequency on the degree to which dialect features are acquired. A feature which mainly applies to frequently used words corresponds to an exemplar cloud which mainly consists of high-frequency items. Each time an item is used it is mapped onto the most similar exemplar, which enhances the lexical strength not only of the relevant exemplar but also of all the exemplars that are associated with it (i.e. all elements of the same cloud). This leads to the activation of an entire pattern (i.e. a dialect feature). So, in the case of a cloud with many high-frequency items (as in the case of a feature applying mainly to frequently used words), the level of activation of all the items is very high, resulting in considerable lexical strength. This explains why dialect features with a high average token frequency are acquired better.

Although a rule-based model cannot account for the fact that frequently used words are more often realized correctly than infrequently used words, the positive effect of average token frequency can more or less be explained within a rule-based framework as well. Actually, the account is comparable to the one concerning the positive effect of incidence (see section 8.2.1). The account runs as follows: If a dialect feature mainly occurs in frequently used words, language users are more often exposed to the relevant feature than if a dialect feature mainly occurs in infrequent words. From a rule-based perspective, we can argue that there is a positive relationship between the degree of exposure to a feature – either because the feature has a high incidence (see section 8.2.1) or a high average frequency – and the ease with which dialect learners form a rule with respect to the relevant feature. Furthermore, a rule-based model predicts that once the learner has formed a rule, the degree of knowledge/acquisition of the relevant feature will increase.

Since there was a significant interaction between average frequency and home language (see table 7.8), in this case we can draw conclusions about the effect of this variable on the degree of dialect knowledge of the native dialect speakers compared to the degree of dialect acquisition of the second dialect learners. It turned out that the positive effect of average frequency is stronger in the case of the native speakers than in the case of the control group. Similar results were found for the effects of incidence and number of SD variants. These

results seem to indicate that first and second dialect acquisition generally follow the same direction.

8.2.5. Accounting for the effects of productivity

In section 4.2.6, we defined productivity (of dialect features) as the property of being applicable to newly introduced words, such as loanwords, proper names or brand names. Our hypothesis with respect to productivity stated that productive dialect features are acquired better than unproductive ones. We therefore expected a positive effect of productivity on the degree to which dialect features are acquired. In section 7.2.2.5 (see table 7.9, figure 7.7), we showed that there is indeed a highly significant positive effect of productivity on the degree of dialect acquisition of the second dialect learners and the degree of dialect proficiency of the control group. There also appeared to be a significant interaction with home language, in that the positive effect of productivity was stronger for the native dialect speakers than for the second dialect learners. Recall that incidence, number of Standard Dutch variants, and average token frequency also had stronger effects in the case of the control group. Like productivity (which distinguishes the postlexical from the lexical features), these are all factors which are supposed to contribute to the degree of predictability of dialect features. As mentioned before, these results indicate that the factors that guide second dialect acquisition parallel those that guide first dialect acquisition.

Recall that we have argued that in the present study, only the postlexical features/rules are productive (see chapter 6, sections 6.4.1-6.4.3). So, the significant effect of productivity on dialect acquisition may also indicate an effect of the distinction between lexical and postlexical features, as well as an effect of other characteristics which are related to this distinction (e.g. ± exceptionlessness). We should therefore be careful in attributing the observed positive effect completely and solely to the property of productivity. The effect could as well be due to other characteristics of postlexical features as compared to lexical ones.[6] In particular, the automatic character and exceptionlessness of postlexical features may be properties that have a favourable effect on the degree to which these features are acquired.[7] It seems plausible that the fact that postlexical features do not allow lexical exceptions makes these features more learnable: the feature applies whenever the structural conditions are satisfied and it is impossible to overapply it (i.e. to make overgeneralization errors).

Apart from the other properties of postlexical features, productivity itself most likely contributes to the positive effect on dialect acquisition as well. A possible line of reasoning could be that productivity directly leads to a high(er) incidence: since productive dialect

[6] The other characteristics of the lexical features, which are responsible for the degree of dialect knowledge/acquisition, were made explicit in the forward regression analyses in section 7.2.2.8.

[7] Recall from section 3.2.3.1 that we argued that the claim that postlexical rules are always exceptionless has been refuted in the literature (cf. Taeldeman 2006b). The postlexical rules involved in this study, however, are all exceptionless rules.

features still apply to new words, they automatically apply to a very large number of words. Actually, the number of words which the postlexical features apply to, could not be calculated because, in principle, this is an 'infinite' number (see section 5.2.1.2).

At first sight, a factor such as productivity, which is closely related to the distinction lexical vs. postlexical, seems to fit better in a rule-based model than in an exemplar-based model. For example, a (postlexical) rule changing all words at once (cf. the 'exceptionlessness' of Neogrammarian rules), without lexical exceptions, can be better interpreted in the context of a model that assumes an underlying, phonological level than in a model assuming word-by-word storage.[8] However, the effect of productivity on dialect acquisition does not exclude an exemplar-based account. One possible explanation may be found in the fact that the productive features involved are all exceptionless features. Recall that we have argued that an exemplar-based model predicts that acquisition is easier in the case of homogeneous exemplar clouds. Exceptionless features (in this case postlexical features, e.g. *n*-deletion; see section 6.4.2) typically involve homogeneous clouds, in which all words with sound(s) X refer to type Y (e.g. all words with 'V + *n* + *s/z*' refer to the dialect variant with 'V: + *s/z*'). Another way in which an exemplar-based model may account for the effect of productivity is by referring to the fact that productive dialect features apply to an infinite number of words and may therefore be acquired more successfully (this is comparable to the positive effect of incidence on acquisition; see section 8.2.1 for the exemplar-based account of the effects of incidence).

8.2.6. Accounting for the effects of geographical distribution

In section 7.2.2.6, we showed that the geographical distribution of dialect features has a highly significant, negative effect on the degree to which dialect features are acquired. This result confirmed our hypothesis that features with a small(er) geographical distribution are acquired most successfully (see section 4.2.7). Recall that this hypothesis was supported by findings from the literature: Hinskens (1986) as well as Auer (1993) observed that when a dialect is acquired as an 'exogenous' system, it is the primary features of a dialect (i.e. the most local and most salient ones; cf. Schirmunski 1930) that are acquired first.

We believe that the observed negative effect of geographical distribution on dialect acquisition should not be interpreted solely in terms of geographical distribution, but also in terms of other, psycholinguistic factors. One of the reasons to assume that the degree of geographical spread cannot be the sole factor responsible for the negative effect, is the fact that children are not supposed to have a very good knowledge/awareness of the geographical distribution of dialect features. The second dialect learners involved in this study may have a

[8] Note that exemplar theory is primarily a theory of language acquisition, and not of historical processes of language change.

slightly stronger awareness of geographical distribution, since most of them have relatives from outside the research location (i.e. one or both of their parents were born and raised elsewhere), so that they can compare the dialect spoken in the research location (i.e. the Maldegem dialect) with the dialect(s) spoken by their family members. However, native dialect speakers do not normally have a good knowledge of the geographical distribution of dialect features of their own dialect. So, the fact that there is also a significant negative effect of geographical distribution on the acquisition of the native dialect speakers seems to suggest that other factors are responsible, because the effect is stronger yet in the case of the native dialect speakers (see figure 7.8).

We have argued, for example, that there is most likely a close relationship between the geographical distribution of a dialect feature and its degree of (perceptual) salience. It was already noted by Schirmunski (1930) that the most local dialect features (i.e. the primary features) are generally the most salient ones (see also Hinskens 1986; Taeldeman 1993, 2006a). Evidence comes, for example, from shibboleths (see De Tier & Vandekerckhove 2003). Shibboleths are used to mock the most typical features of a particular dialect (mostly by speakers of neighbouring dialects which do not have the feature(s) that are made fun of). This seems to indicate that it is the most local features which are most salient, especially to people speaking another dialect. An example of a shibboleth concerning the Maldegem dialect is the phrase *vijfenvijftig ijzeren vijzen* 'fifty-five iron screws', which refers to the primary feature SD /ɛi/ ~ DIA /ø/. The fact that this feature occurs in a shibboleth indicates that it is a very salient feature of the Maldegem dialect. In addition, it has a very small geographical distribution (= 1; see appendix 4).

Whether there is a negative relationship between the degree of geographical distribution of the features involved, and the degree of salience (i.e. whether the smallest geographical distribution corresponds to the highest degree of salience and vice versa), has to be investigated independently in further research. So, we can only account for the negative effect of geographical distribution by hypothesizing that the most salient features are acquired most successfully.

A possible explanation why some dialect features are more salient than others might be found in the phonetic distance between Maldegem dialect elements on the one hand, and equivalent elements in the surrounding dialects and Standard Dutch on the other (see section 4.2.7). Taeldeman (1993) suggests, for example, that secondary dialect features are characterized by a large geographical distribution and a small phonetic distance towards the other variety involved. Furthermore, 'phonetic distance' is part of Trudgill's (1986) 'definition' of salience. We have pointed out that several methods to measure phonetic distance have been proposed (cf. Cucchiarini 1993; Hoppenbrouwers & Hoppenbrouwers 1988; Heeringa 2004). It would be interesting to examine whether the Maldegem dialect features with the smallest geographical distribution are the ones with the largest phonetic distance towards equivalent elements from surrounding dialects or Standard Dutch. However,

this would be beyond the scope of the present research. We presume that the phonetic distance between Maldegem dialect elements on the one hand, and Standard Dutch elements or elements of surrounding dialects on the other, may play a role in determining the success with which features are acquired.

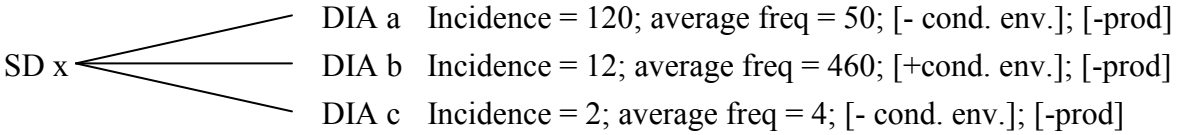
Finally, the question arises why the most typical (and probably the most salient) features of the Maldegem dialect are acquired better both by the second dialect learners and by the control group. This might be due to a strong positive attitude towards the local/target dialect and the place where it is spoken (cf. *Ortsloyalität*, Mattheier 1980). Because of a high degree of *Ortsloyalität*, children will – probably unconsciously – focus mainly on the features which are most salient in the dialect in question (i.e. most salient as opposed to the surrounding dialects and the standard language) (see also Taeldeman 2006a).

Hinskens (1986) and Auer (1993) claimed that the primary features of a dialect are acquired *first*, when the relevant dialect is acquired as an exogenous system (cf. a second language). In order to obtain information about when in the acquisition process particular dialect features are acquired most successfully, we should take into account possible interaction effects with age. The interaction of geographical distribution with age turned out to be significant (see table 7.11 and figure 7.9). Figure 7.9 revealed that the negative effect of geographical distribution on acquisition is somewhat weaker in the case of the fifteen-year-olds, as compared to the nine- and twelve-year-olds. This result indicates that the factor geographical distribution, which is closely related to salience, is more influential to the dialect acquisition of the two younger age groups. This might indicate that the primary dialect features (i.e. those with the smallest geographical distribution) are indeed acquired first, as was suggested by Hinskens (1986) and Auer (1993).

8.2.7. Interactions between factors determining predictability

Recall that we have argued (section 4.2.1) that incidence (for the lexical features), productivity (for the postlexical features), conditioning environment, number of competing dialect (and Standard Dutch) variants, and average token frequency are factors which all contribute to the degree of predictability of dialect features. Recall that we proposed the following hypothetical situation (cf. figure 4.1; section 4.2.1):

Figure 8.1: Factors contributing to the degree of predictability of dialect features



In section 4.2.1 we argued that different factors play a role in the degree of predictability of the three (hypothetical) dialect features represented in 8.1: all features have the value '3' for number of dialect variants, but this does not imply that all three features are equally predictable: the feature SD x ~ DIA a, for example, may gain predictability by its high incidence (see section 8.2.1), whereas the feature SD x ~ DIA b may have an increased degree of predictability because it is restricted to a conditioning environment (see section 8.2.3), or because it has a relatively high average token frequency (see section 8.2.4.2).

As mentioned before, all lexical features in this study are unproductive. The only three postlexical features involved are all productive. However, the factors incidence and average token frequency are not applicable to postlexical features. As a result, there are no significant interactions between productivity and incidence on the one hand, and between productivity and average frequency on the other.

In section 7.2.2.7, we showed that there are many significant interactions between the factors that are supposed to determine the predictability of dialect features. There appeared to be three significant four-way interactions. Furthermore, there were eleven significant three-way interactions involving home language, nine of which were analysed in more detail.[9] From the fact that in many cases there is a significant interaction with home language, it follows that we can draw conclusions about the difference between the second dialect learners and the control group.

Recall that we have argued that only the (significant) highest-order interaction(s) can give a complete picture of the exact nature of the interactions between the different variables. However, we were not able to represent the significant four-way interactions graphically in an insightful manner. However, we did represent most of the significant three-way interactions. It is important to emphasize that we have to be cautious in drawing firm conclusions from these (three-way) interactions, since these are not the highest-order interactions.

In section 8.2.7.1, we summarize the main findings with respect to most of the three-way interactions and, if possible, we try to account for these results. In section 8.2.7.2, we discuss the interaction between average frequency, conditioning environment, and home language in more detail, because we formulated a specific expectation with regard to this interaction (see chapter 4).

8.2.7.1. Possible accounts of the three-way interactions between the factors determining predictability

In table 8.1 we summarize the main findings with respect to the significant three-way interactions between the factors contributing to the degree of predictability of dialect features, and we try to account for these results.

[9] In two cases, it was impracticable to interpret the three-way interactions, because too many diagrams would have been involved (see section 7.2.2.7).

Interaction	Result	Explanation
conditioning environment* n dialect variants* home language	Among the control group (home language = 0), there was a negative effect of number of dialect variants for dialect features without a conditioning environment, which was not found for features with a conditioning environment (in fact, there was an unexpected positive effect, which is very difficult to account for and which cannot be explained here) (see figure 7.10A). For the second dialect learners (home language = 1), the same basic interaction effect was found (see figure 7.10B). The interaction with home language could be seen from the fact that the effects were more pronounced for the second dialect learners (see figure 7.10B): the negative effect for features without a conditioning environment as well as the positive effect for features with a conditioning environment were stronger (as shown by the steeper slopes) in the case of the second dialect learners.	The fact that there is a negative effect of number of dialect variants (for features without a conditioning environment) is understandable: we already argued in section 4.2.2 that a high number of competing forms makes acquisition more difficult (see the relevant section for more details). On the other hand, the presence of conditioning environments means that a dialect learner will be able to find out to a certain extent in what contexts all or at least some of the competing dialect variants should be used. Such indications are not available for paradigms (of competing dialect variants) in which there are no conditioning environments at all. Probably, the fact that learners can benefit from the presence of conditioning environments compensates for the negative effect of a higher number of competing dialect variants.
n dialect variants* n SD variants* home language	Diagrams 7.11A and 7.11B revealed that the best performance occurred when five Standard Dutch variants corresponded to six dialect variants, whereas the worst performance occurred when one Standard Dutch variant corresponded to six dialect variants.	A possible account of this result is that acquisition is easier if the number of competing SD variants and the number of competing dialect variants are well-balanced (e.g. 5 SD variants corresponding to 6 dialect variants) than in the case of an imbalance between the number of SD variants and the number of dialect variants (e.g. 1 SD variant corresponding to 6 dialect variants).

conditioning environment* n SD variants* home language	Among the control group (home language = 0), there was a negative effect of number of SD variants for dialect features without a conditioning environment, which was not found for features with a conditioning environment (actually, there was an unexpected positive effect, which will not be further discussed). For the second dialect learners (home language = 1), a similar interaction effect was found (see figure 7.12B). Again, there is a significant interaction with home language, in the sense that the effects are stronger for the second dialect learners (see figure 7.12B).	The fact that there is a negative effect of number of SD variants (for features without a conditioning environment) could be predicted: section 4.2.2 argued that a high number of competing forms leads to a lower degree of dialect proficiency. We already noted that learners may benefit from the presence of conditioning environments, which compensates for the negative effect of a higher number of competing dialect variants. In this way, the lack of a negative effect of number of SD variants for features with a conditioning environment can be accounted for as well.
conditioning environment* n dialect variants *incidence	From the results presented in figures 7.13A and 7.13B the following generalization can be derived: the presence of a conditioning environment (cf. diagram 7.13A vs. 7.13B) as well as a lower number of competing dialect variants (cf.. diagram 7.13A as well as 7.13B) has a weakening influence on the effect of incidence.	Recall from section 8.2.1 that incidence has a positive effect on the degree to which dialect features are known/acquired. It is not so easy to explain the finding that this effect becomes less strong as a result of the fact that there is a conditioning environment and/or a small number of competing dialect variants. Possibly, it is the case that even though a higher incidence will stimulate learners to give a higher 'priority' to the acquisition of a feature, this effect is weakened if there are other factors which increase its learnability, such as a conditioning environment or a small(er) number of competing dialect variants. Note that in some cases incidence even has a negative effect. We do not have an account for this result, which is in conflict with the general picture with respect to the effect of incidence (see section 8.2.1).
conditioning environment* n SD variants* incidence	The interaction between number of SD variants and incidence was only significant for dialect features with a conditioning environment (see diagram 7.14). Diagram 7.14 showed that the positive basic effect that was found for incidence (see section 8.2.1) does not apply to features involving a low(er) number of SD variants.	Presumably, this result can be related to the same consideration as made above, i.e. that even though a higher incidence will trigger learners to give a higher 'priority' to the acquisition of a feature, this effect is weakened if there are other factors which increase its learnability, in this case a small(er) number of competing SD variants.

conditioning environment* n dialect variants* n SD variants	For features without a conditioning environment (see figure 7.15A), there was an overall negative effect of n dialect variants, which was strongest in the case of features involving the maximum number of competing SD variants. The same effects could be seen for features with a conditioning environment (see figure 7.15B), except that n dialect variants had an (unexpected) positive effect for features with 4 or 5 SD variants.	The interaction effects of n dialect variants with n SD variants for features without a conditioning environment (see figure 7.15A) are quite plausible: acquisition is most difficult if there are 6 competing dialect variants and 5 competing SD variants, whereas acquisition is most straightforward if there are only two competing dialect variants and only one SD variant. In the case of features with a conditioning environment (see figure 7.15B), we cannot explain why there is a positive effect of n dialect variants for features involving 4 or 5 SD variants.
average token frequency* n dialect variants *home language	Diagrams 7.16A and 7.16B revealed that the positive effect of average token frequency on dialect acquisition is strongest in the case of dialect features involving the largest number of competing dialect variants (i.e. 6 dialect variants), whereas the effect is weakest in the case of features involving the smallest number of competing dialect variants (i.e. 2 dialect variants). These interaction effects were comparable for the native dialect speakers and the second dialect learners.	Again, this result can be related to the same consideration as those made with respect to incidence, i.e. that even though a higher average frequency will stimulate learners to give a higher 'priority' to the acquisition of a feature, this effect is weakened if there are other factors which increase its learnability, in this case a small(er) number of competing dialect variants.
productivity* n SD variants* home language	The interaction between n SD variants and productivity was only significant in the case of the native dialect speakers (home language = 0) (see figure 7.17). The effect of number of SD variants was slightly negative for unproductive features (in this case: the lexical features), whereas it was positive in the case of productive features (in this case, the postlexical features).	This result may indicate that the negative effect of number of SD variants on the degree of dialect acquisition becomes less important if there is a factor which enhances the learnability of features, such as productivity, which has a positive main effect on the degree of acquisition (see section 8.2.5).

Table 8.1: Main findings with respect to the interaction effects between the factors contributing to the degree of predictability and possible accounts

Summarizing, we can conclude that the interaction effects described in table 8.1 are sometimes difficult to account for, but there seems to be one general tendency: if a certain factor has a positive effect on the degree of dialect knowledge/acquisition, this effect becomes weaker in interaction with other factors which also enhance the degree of dialect knowledge/acquisition.

In the next section, we give a more detailed discussion of the three-way interaction between average frequency, conditioning environment, and home language, which was not described in table 8.1.

8.2.7.2. Average token frequency*conditioning environment*home language

In section 7.2.2.7.9, we showed that there is a significant interaction effect between average token frequency, conditioning environment, and home language. Beforehand, we had certain expectations about this interaction: in section 4.2.4, we suggested that there is probably an interaction between conditioning environment and average frequency. Phonological neighbourhoods which mainly consist of frequently used words or word forms (i.e. with a high average frequency) were predicted to have a stronger positive effect on the degree of dialect knowledge/acquisition than neighbourhoods which mainly consist of infrequently used words or word forms.

Figure 7.18A represented the interaction between average token frequency and conditioning environment for the group of native dialect speakers (home language = 0), and figure 7.18B visualized this interaction for the group of second dialect learners. Among the control group (fig. 7.18A), there was a positive effect of average token frequency for dialect features without a conditioning environment, whereas there was a very weak negative effect for dialect features with a conditioning environment. So, as far as the native dialect speakers are concerned, the exemplar-based prediction that the (positive) effect of conditioning environment would be stronger if the average frequency is high, is not fulfilled.

The interaction with home language is clear from a comparison of diagrams 7.18A and 7.18B: for the second dialect learners (see fig. 7.18B), the effect of average token frequency is positive, both in the case of features without a conditioning environment and in the case of features with a conditioning environment. In fact, the positive effect of average frequency is even stronger in the case of dialect features that are not restricted to a conditioning environment.

On the basis of diagrams 7.18A and 7.18B we can therefore conclude that average token frequency has a stronger (positive) effect when there is no conditioning environment. This result is consistent with earlier findings: in the accounts given in table 8.1, we argued that the presence of a conditioning environment may compensate for the negative effects of number of dialect variants and number of SD variants, respectively. Furthermore, we suggested that the positive effect of incidence and of average token frequency might be weakened if there are other factors which increase learnability (such as a low number of competing dialect variants). Likewise, the results in figures 7.18A and 7.18B can be explained as follows: the positive effect of average token frequency on the degree of knowledge or acquisition of a particular dialect feature may be weakened if there are other factors which increase the learnability of the feature, such as the presence of a conditioning environment. On the other hand, the effect

of average frequency is enhanced if there are no other factors (such as a conditioning environment) which make acquisition easier.

Summarizing, we may conclude that the exemplar-based prediction that phonological neighbourhood effects are stronger when the relevant neighbourhoods mainly contain frequently used words, is not borne out as far as the native dialect speakers are concerned (cf. figure 7.18A), whereas this prediction is confirmed for the second dialect learners (cf. figure 7.18B). However, in the latter case, the effect of average frequency is stronger for features without a conditioning environment.

8.2.8. Summary

In this section, we have accounted for the effects of the feature-related factors on the degree to which dialect features are acquired. Most of these effects could be accounted for both within a rule-based framework and within an exemplar-based framework. Only with respect to the factor token frequency (on the level of the word), we may conclude that an exemplar-based model is more suitable. The fact that token frequency has a significant, positive effect on the degree of dialect acquisition (of individual words) cannot easily be accounted for in a model which assumes that there is a separate module of phonological representation which is independent of word-specific information (i.e. a rule-based model). On the other hand, exemplar-based models offer a satisfactory account of frequency effects, because these models assume that learners have direct access to word-specific information (i.e. words and their idiosyncratic details are individually memorized).

Most factors displayed similar effects for the second dialect learners and the native dialect speakers. However, generally these effects were stronger in the case of the native dialect speakers. This was the case for incidence, average token frequency, productivity, and geographical distribution. These results indicate that even the native dialect speakers have no 'perfect' knowledge of the Maldegem dialect (yet), and that they are probably still in a process of dialect acquisition, just like the second dialect learners. However, they have made more progress than the second dialect learners (see table 7.1, section 7.2.1), which does not come as a surprise, since they started earlier. Our results reveal that the factors that guide second dialect acquisition parallel those that guide the acquisition of a dialect as a first language.

One of the factors which we supposed would play a role in the process of (second) dialect acquisition was the number of competing dialect variants, but this factor did not turn out to have a significant effect (see sections 7.2.2.2.1 and 8.2.2.1). However, there appeared to be significant interactions between number of dialect variants and some of the other factors which were hypothesized to contribute to the predictability of dialect features (i.e. incidence, n SD variants, conditioning environment, productivity, and average frequency). Furthermore, the factor number of dialect variants also had a significant effect on the degree of dialect

knowledge when entered in stepwise regression analyses involving the other feature-related factors as well (except for productivity) (see table 7.21A and 7.21B, section 7.2.2.8).

A comparison of the most important factors determining the degree of dialect knowledge for the native dialect speakers and for the second dialect learners, respectively, indicated that there is a high degree of similarity between both groups: in both cases, incidence had the largest relative effect, followed by conditioning environment and number of dialect variants. This result also indicates that second dialect acquisition and first dialect acquisition generally follow the same direction as far as the acquisition of the lexical features is concerned (productivity could not be included in the stepwise regression analyses, so that the postlexical features were excluded).

Only the factor geographical distribution significantly interacted with age. The interaction with age revealed the relative importance of geographical distribution in different stages of the process of dialect acquisition. We found that the effect of geographical distribution (related to salience, cf. above) becomes weaker as children grow older (i.e. a weaker effect was found for the fifteen year-olds). The fact that no further interactions with age were found, indicates that the effects of the other feature-related factors do not significantly depend on the age group to which children belong. It could be the case that more significant interactions with age would have been found if we had examined the speech of children younger than nine or older than fifteen.

8.3. Discussing the effects of speaker-related factors on the individual success of (second) dialect acquisition

8.3.1. Accounting for the effects of age

In section 7.2.3.1, we showed that the degree of dialect proficiency continues to increase between the ages of nine and fifteen. No significant interaction of age with home language was found, despite the fact that we would have expected that the native speakers reached their ceiling sooner than the dialect learners. Although the native speakers continue to improve their dialect proficiency at the same rate as the second dialect learners, they reach a higher level of dialect proficiency than the dialect learners when they are nine years old, and this higher level of dialect proficiency continues to exist at the ages of twelve and fifteen. The fact that the native speakers do not reach their ceiling before the age of fifteen can only be accounted for by assuming that these children are still in a process of dialect acquisition. This may be related to the fact that it takes longer to acquire a dialect 'perfectly' if that dialect has been exposed to processes of dialect levelling.

Our data did not reveal that there is a critical age of (first/second) dialect acquisition, after which acquisition slows down. We cannot exclude the possibility that there is a critical age of dialect acquisition after the age of fifteen.

8.3.2. Accounting for the effects of gender

In section 4.3.2, we pointed out that many studies in language variation and change have observed a tendency of females to speak less dialectically than males (cf. Chambers & Trudgill 1980; Cheshire 1982; Hoppenbrouwers 1990; Taeldeman 1991, 1995). Females seem to be more sensitive to processes of dialect levelling. Following this line of reasoning, it could be expected that gender causes variation *among the native speakers* of the Maldegem dialect, in that the boys will speak more dialectically than the girls. Under the assumption that there is indeed a difference between boys and girls among the control group, it might be expected that gender has an effect on the degree of dialect acquisition of the second dialect learners as well: male dialect learners might make more efforts than females to acquire the local dialect, because the male native speakers speak more dialectically than the female native speakers.

However, in section 7.2.3.2 we showed that there is no significant effect of gender on the degree of dialect knowledge among the native dialect speakers (see table 7.23). This implies that the line of reasoning described above does not hold true for our data. There is no evidence that the native dialect speakers of the Maldegem dialect are – as in many other places in Flanders – gradually moving into the direction of Standard Dutch, with the girls taking the lead. This implies that the degree to which the male native speakers speak the local dialect does not significantly differ from that of the female native speakers. Therefore, the fact that we found a significant effect of gender for the second dialect learners (with boys being more successful than girls), cannot be due to differences among the native Maldegem male and female peers.

Our finding that the boys from the group of second dialect learners are more successful in dialect acquisition than the girls from this group, is not in agreement with Vousten's results. Recall from section 4.3.2 that Vousten (1995) did not find a significant difference between the dialect learning boys and the dialect learning girls. Vousten accounted for this result as follows. In processes of second dialect acquisition, we are dealing with standard language speakers who acquire a dialect. Acquiring the dialect does not mean, however, that children have to give up their L1 (= Standard Dutch) in favour of that dialect: they just add a dialectal variety to their linguistic repertoire. The fact that girls generally attach more value to the prestige of the standard variety is therefore not relevant in the context of second dialect acquisition. However, our results may indicate that girls who acquire a dialect as a second language are more reserved with the use of dialect forms than boys. This may be related to the fact that girls are more sensitive to factors of prestige (cf. Chambers & Trudgill 1980). Furthermore, Taeldeman (1991:section 3) argued that boys make more efforts to acquire the local dialect through channels other than the family (e.g. through friends at school or in the youth movement) than girls.

8.3.3. Accounting for the effects of attitude/motivation

We have proposed the hypothesis (see section 4.3.3) that children with a positive attitude towards the Maldegem dialect and dialect use/speakers and with a strong motivation to learn the local dialect will be more successful in (second) dialect acquisition than children with a negative attitude. This hypothesis was confirmed by the results (see section 7.2.3.3): there is a highly significant, positive effect of attitude on the degree to which the correct dialect variant is realized, both in the case of the second dialect learners and of the control group. Since there was no significant interaction with home language, however, we cannot draw conclusions about the effect of attitude that are specific for second dialect acquisition alone. The (positive) effect of attitude is as strong for native dialect speakers as for the second dialect learners.

Before interpreting this result, it is important to note that there was a significant difference with respect to the score for attitude between the native dialect speakers and the second dialect learners. The control group had an average score for attitude/motivation of 9.35%, whereas the second dialect learners had an average score of 7.52%. The difference between these scores appeared to be significant ($F = 519.2$, $df = 24325$, $p = .000$). The fact that the second dialect learners have a less positive attitude towards the local dialect and dialect use/speakers might be related to the fact that these children are likely to have 'inherited' this attitude from their parents, who probably have less positive attitudes towards dialect use themselves (otherwise, one might argue, they would have raised their children in the local dialect).

At first sight, the fact that attitude has an equally strong effect on the degree of dialect knowledge of the native speakers as on the degree of knowledge of the dialect learners, may be surprising. We might have expected that attitude had a weaker effect on the degree of success in first dialect acquisition than on the degree of success in second dialect acquisition, since first dialect acquisition normally proceeds in a less conscious way. Furthermore, speakers need their mother tongue to communicate anyway. If a child acquires a dialect as a second language (at an older age), it can 'afford' to have a (passive or active) negative attitude towards dialect use. However, such a result was not found. Probably, this can be related to the fact that (a number of) members of the control group had not acquired a profound knowledge of the Maldegem dialect at the time of the interviews. This would imply that for these children, the task of dialect acquisition as a first language is a more conscious process than expected (because it is only accomplished at a later point in time), which would explain that the predicted contrast between members of the control group and second dialect learners was not confirmed.

One way to account for the positive effect of attitude/motivation on the degree of dialect knowledge is by assuming that children with a positive attitude towards the Maldegem dialect will make more linguistic efforts to integrate into the dialect speaking peer group. However, it might well be the other way around: children who have many friends in the dialect speaking

peer group and who are very peer group-oriented, may develop a more positive attitude towards the Maldegem dialect than more isolated children or children who have their friends mainly outside the school. The fact that there is an equally strong effect of attitude on the dialect proficiency of the native dialect speakers as on the degree of dialect proficiency of the second dialect learners, may also result from the fact that even among the native dialect speakers there are children whose friends mainly speak Standard Dutch or a substandard variety (in this case *tussentaal*), affecting their own attitude towards dialect use and interfering with their dialect proficiency.[10,11] Summarizing, it is plausible that attitude is a factor that depends to a large extent on the child's position within the group. It would therefore be interesting to do more research into the child's preferences within the group, for example by means of a sociogram (cf. Berthele 2000, see chapter 2).

8.3.4. Accounting for the effects of origin of the parents

Recall that we hypothesized (see section 4.3.4) that children (raised in Standard Dutch or *tussentaal*) with one or both parents born in Maldegem acquire the Maldegem dialect more successfully than children with one or both parents from elsewhere. This implies that we expected a positive effect of the origin of the mother and of the father on the degree of success in dialect acquisition.[12] Further, we had reasons to assume (cf. Taeldeman 1986; Labov 1990; Roberts 2001) that the effect of the origin of the mother is more decisive than the effect of the origin of the father.

Unexpectedly, the results (see section 7.2.3.4) revealed a highly significant, *negative* effect of the origin of the mother on dialect acquisition, indicating that children whose mother came from Maldegem acquired the Maldegem dialect less successfully than children whose mother came from elsewhere (see table 7.26). As expected, the effect of the origin of the father was positive (see table 7.27), which means that children with a father from Maldegem acquired the local dialect more successfully than children with a father from elsewhere. Finally, also contrary to the hypothesis, the effect of the origin of the father appeared to have most relative weight in a model that contains the variables 'origin mother' and 'origin father' (see table 7.28).

The main question raised by these results is: how can the effect of the origin of the mother be negative? We propose the following hypothesis to account for this result. From the literature (cf. Chambers & Trudgill 1980; Hoppenbrouwers 1990; Taeldeman 1991) it appears

[10] Recall that we have defined *tussentaal* as a "language variety that makes use of dialectal elements, while at the same time operating on a supraregional level, as a direct competitor to Standard Dutch" (Plevoets et al. 2006) (see section 6.3 for a more detailed discussion of *tussentaal*).
[11] There may also be a discrepancy between a child's knowledge of the local dialect and its willingness to speak that dialect. The unwillingness to speak dialect may be related to the fact that dialect is associated with a lower degree of social prestige (cf. the notion of overt prestige, see chapter 2) in Flanders (cf. Taeldeman 1991).
[12] Recall that we coded origin of the mother/father as "1" if the mother/father was born in Maldegem, and as "0" if the mother/father came from elsewhere.

that female language users are inclined to use more standard(ized) forms (and thus, fewer dialectal forms) than males. As in most other Flemish places, this tendency can be seen in Maldegem as well (cf. Versieck 1989:190).[13] It might be the case, then, that the Maldegem mothers speak less dialectically with their children than the Maldegem fathers. Moreover, even mothers who were born in Maldegem are not necessarily native speakers of the Maldegem dialect.

Generally, the reasons for females to speak less dialectically are partly related to attitudinal factors (cf. Chambers & Trudgill 1980; Lebbe 1996:158-159).[14] It seems plausible that mothers have more pronounced (negative) attitudes towards their child speaking the local dialect than fathers. This might be related to the fact that mothers are more closely involved in the linguistic development of their children than fathers. However, this still does not answer the question why children with mothers from Maldegem are generally less successful dialect learners than children with mothers from elsewhere. We believe that a possible answer to this question may be that mothers who are from Maldegem and who speak the Maldegem dialect themselves, think in a more negative way about this dialect than mothers who are not from Maldegem and who do not speak the local dialect. "Linguistic self-hatred" (Labov 1966:489) is not untypical of dialect speakers. It might be the case that stronger negative attitudes are involved vis-à-vis the own language variety than in the case of an exogenous variety (e.g. the Maldegem mothers might consider the local dialect as a 'threat' for their children's proficiency in the standard language). On the other hand, mothers who are not from Maldegem and therefore do not speak the local dialect themselves, may have less pronounced negative attitudes towards dialect use than the mothers who are from Maldegem. This may be related to the fact that mothers from elsewhere (i.e. mothers who do not belong to the proper Maldegem dialect community themselves) are more enthusiastic about the fact that their children integrate in the dialect speaking peer group by learning the local dialect.

As expected, the origin of the father appeared to have a *positive* effect (see table 7.27) on the degree to which dialect features were realized correctly. This means that children (raised in Standard Dutch or *tussentaal*) whose father comes from Maldegem acquire the Maldegem dialect more successfully than children whose father comes from elsewhere. This result possibly indicates that fathers are more inclined to switch or keep to the Maldegem dialect in home situations, despite their intention to raise their children in a non-dialectal variety. This may be related to the fact that fathers generally are less concerned about the linguistic development of their children than mothers. In addition, males have less negative attitudes towards dialect use and speak more dialectically than females (cf. Münstermann & Van Hout 1988).

[13] The tendency of females to speak less dialectically than males was not found in our data with respect to native dialect speakers between nine and fifteen years old, however.
[14] Actually, the differences between males and females are considerably larger with respect to attitudes than with respect to actual dialect use (cf. Jacobs 1990).

Note that, absolutely speaking, the origin of the father had a larger effect on the degree of dialect acquisition than the origin of the mother (see table 7.28). This implies that the positive effect of having a Maldegem father is more decisive in second dialect acquisition than the negative effect of having a Maldegem mother. This result contradicts the findings from the literature (cf. Taeldeman 1986, Roberts 2001): apparently, the fact that the mother is often the most important caretaker in the earliest years of life does not cause complete orientation towards the mother's language later in life, as far as the second dialect learners are concerned. A possible reason is that the predominance of the mother decreases as the child grows older. Probably, we would have found a more important effect of the mother if we had examined younger children.

8.3.5. Summary

In this section we have shown that speaker-related factors such as age, gender, attitude and origin of the parents have significant effects on the degree of dialect knowledge/acquisition of individual language users.

We have shown that the degree of dialect proficiency keeps increasing between the ages of nine and fifteen, and that there are no indications for a critical age of language learning during this period. It was argued that native speakers are still in a process of dialect acquisition themselves, which explains why their degree of dialect knowledge continues to improve at the same rate as that of the second dialect learners.

We have shown that there is a significant effect of gender on the degree of second dialect acquisition, in that the boys are more successful learners than the girls. We have argued that a possible explanation for this result can be found in the fact that girls attach more importance to factors of prestige, whereas boys generally make more efforts to acquire the local dialect through channels other than the family.

There was also a significant positive effect of attitude on the degree of dialect knowledge. We have argued that the attitude of a child towards the local dialect and dialect use probably depends on the child's position in the peer group.

Finally, we have proposed a hypothetical account of the negative effect of the origin of the mother on dialect proficiency. Mothers from the research location (most of whom had a good to very good knowledge of the Maldegem dialect themselves) may have stronger negative attitudes towards the Maldegem dialect than mothers from elsewhere.[15] The effect of the origin of the father was positive, probably because fathers are more inclined to switch or keep to their own dialect, even if they have the intention of raising their children in Standard Dutch.

[15] In the written questionnaire that had to be filled in by the parents, we inquired into the dialect knowledge of the parents.

Recall from the stepwise regression analyses in section 7.2.3.5 that age is by far the most important speaker-related factor in determining the degree of dialect knowledge of native dialect speakers and second dialect learners. We may therefore conclude that the period of exposure to the Maldegem dialect (i.e. the period of nurture), but probably also the (age-related) degree of peer group orientation (recall the fact that attitude is the second most important factor determining the degree of dialect knowledge) and the adaptability of the innate language-learning device (i.e. 'nature') are – apart from feature-related factors – most decisive for the success in dialect proficiency. Again, it appears that first and second dialect acquisition mainly follow the same direction.

8.4. Discussing the effects of feature-related factors on the production of overgeneralizations

8.4.1. Introduction

In section 7.3.1, we gave an overview of the degree to which the native dialect speakers and the second dialect learners produced overgeneralization errors of different types (see table 7.30). There appeared to be a significant overall difference between the degree to which the native speakers produced overgeneralizations (within the /ɛi/-paradigm) and the degree to which the dialect learners made overgeneralization errors.

In almost all cases where there was a significant difference between both groups, the second dialect learners made more overgeneralization errors than the control group. An exception was the overgeneralization of SD /a:/ → DIA /ɔ̟ᶜ/ at the expense of SD /a:/ → DIA /ɑ/ in the words *schaatsen* 'to skate' and *laatste* 'last one'. In that case, the control group made significantly more overgeneralization errors than the dialect learners. As we pointed out in section 7.2.1, this overgeneralization has become so common among native dialect speakers that it has more or less become the 'new' dialect variant. This explains the high number of 'overgeneralizations' in the dialect of the control group. This raises the question why the second dialect learners produce this type of overgeneralization less frequently. We hypothesize that there may be two reasons: first, the second dialect learners may fall back on their first language (i.e. Standard Dutch or the substandard) and thus produce another form than the 'overgeneralized' form. Second, table 7.1 revealed that the second dialect learners were more successful in producing the correct dialect variant in the words *schaatsen* and *laatste*. We assume that this is due to the fact that they guessed the correct form 'by coincidence'.

Although table 7.30 revealed that in the overwhelming majority of cases the native dialect speakers produce significantly fewer overgeneralizations than the second dialect learners, there was still a considerable number of overgeneralizations in the dialect of the native

speakers. This indicates that – as we have argued before – the native dialect speakers were still in a process of dialect acquisition themselves and had not yet achieved a 'perfect' knowledge of the Maldegem dialect. It does not come as a surprise that overgeneralization errors are not only typical of second dialect acquisition, but are also part of first dialect acquisition.

In section 7.3.2, we discussed the results of the effects of a number of feature-related factors on the degree to which dialect features are overgeneralized. Recall that we argued that there is not necessarily a negative correlation between the degree to which a dialect feature is known (i.e. realized correctly) by our informants, and the degree to which that feature is overgeneralized (see section 7.3.2.1). A particular feature may score 100% for acquisition, and at the same time be overgeneralized to a large extent. This implies that a score of 100% in our study does not mean that perfect acquisition has been achieved. It only implies that a particular informant realized all words to which the feature applies correctly. This does not exclude the possibility that this informant overgeneralized that feature in words where another feature should have been used. There is not necessarily a positive correlation between the degree to which a feature is known and the degree to which it is overgeneralized either. So, the factors which affected the degree of dialect knowledge/acquisition do not necessarily have parallel effects on the degree of learning and on the degree of overgeneralization.

In the following sections, we will try to account for the effects of the feature-related factors on the degree of overgeneralization.

8.4.2. Accounting for the effects of incidence

The results discussed in section 7.3.2.2 indicated that the incidence of dialect features has a highly significant, positive effect on the degree to which these features are overgeneralized. This means that a feature such as SD /ɛi/ → DIA /ø/, with an incidence of 40, is overgeneralized more often than a feature such as SD /ɛi/ → DIA /æi/, with an incidence of 8. Moreover, there was a significant interaction with home language, in that the positive effect of incidence was stronger in the case of the second dialect learners than in the case of the control group.

The positive effect of incidence on the degree of overgeneralization was to be expected. In section 4.4.4.2, we formulated the hypothesis that a higher incidence would lead to a larger number of overgeneralization errors. We have already argued why this hypothesis is defensible from a rule-based as well as an exemplar-based perspective. Below, we briefly recapitulate the discussion presented in section 4.4.4.2.

A rule-based account of the positive effect of incidence may run as follows. Rule-based theory considers overgeneralizations as the result of the overapplication of a rule. The assumption is that a (correspondence) rule has been formed first. It is easier for a child to

form a (correspondence) rule that generalizes over a particular dialect feature, if that feature occurs in a large number of words, i.e. if the feature has a high incidence. Because of the high incidence, language users are frequently exposed to the relevant feature, so that the rule that generalizes over this feature will be more readily overgeneralized (at the expense of other features), as long as the feature is not acquired 'perfectly'.

From an exemplar-based perspective, the positive effect of incidence on overgeneralization may be accounted for as follows. Dialect features with a high incidence correspond to a large exemplar cloud. The chance that an exemplar is activated within a large cloud is higher than the chance that an exemplar belonging to a small cloud is activated. The activation of one item results in the activation of all associated exemplars (i.e. its nearest neighbours), so that the level of activation of the entire cloud (i.e. and also of the dialect feature) becomes very high. So, exemplar clouds that correspond to features with a high incidence are usually stronger than those which correspond to features with a low incidence. We have also argued that new words (i.e. words of which the learner does not know the correct dialect variant yet) are attracted to their nearest neighbours (i.e. the dialect variant of the nearest neighbour is directly copied onto the new word), especially to the nearest neighbours that belong to a strong exemplar cloud. However, this process does not always result in the correct dialect variant (of the new word); it also leads to overgeneralization errors, i.e. when the wrong dialect variant is copied onto the new word.

Summarizing, we can conclude that both models of language learning (rule-based and exemplar-based) provide a satisfactory account of the positive effect of incidence on the degree of overgeneralization.

The fact that incidence has a similar effect on the degree of overgeneralization among the second dialect learners as among the native speakers, indicates (again) that the factors that guide second dialect acquisition act in the same way in the process of first dialect acquisition. The effect of incidence on the degree of overgeneralization appeared to be strongest for the second dialect learners.

8.4.3. Accounting for the effects of conditioning environment

In section 7.3.2.3, we showed that the factor conditioning environment has a significant positive effect on the degree to which dialect features are overgeneralized. This indicates that features with a conditioning environment (e.g. the feature SD /ɛi/ → DIA /ø/, before anterior consonant) are overgeneralized more frequently than features without a conditioning environment (e.g. SD /ɛi/ → DIA /æi/, lexically determined). There was no significant interaction with home language, indicating that the effect of conditioning environment on the degree of overgeneralization is equally strong for the second dialect learners as for the native dialect speakers.

A positive effect of conditioning environment on the degree of overgeneralization was predicted by an exemplar-based model (see section 4.4.4.3). Such a model accounts for this positive effect as follows. Dialect feature A (e.g. SD /ɛi/ → DIA /e/), which is restricted to a conditioning environment (e.g. before velar/laryngeal consonant), very likely corresponds to a phonological neighbourhood (i.e. an exemplar cloud formed on the basis of phonological similarity) which is formed on the basis of the relevant conditioning environment (e.g. a neighbourhood containing words like *kijken* 'to look', *rijk* 'rich', *krijgen* 'to receive', etc.). New words that should have dialect feature A (e.g. *lijk* 'corpse', *vijg* 'fig') are likely to be attracted to the nearest neighbour in the neighbourhood that corresponds to dialect feature A, and as a result, these words receive the correct dialect variant (this explains the positive effect of conditioning environment on the degree to which dialect features were realized correctly; see section 8.2.3). On the other hand, in the case of dialect feature B (e.g. SD /ɛi/ → DIA /æi/), which has no conditioning environment, the question whether a particular word has dialect feature B or not is much more arbitrary (i.e. there is no environmental conditioning). As a result, words that should have dialect feature B (e.g. *eigen* 'own', *reiger* 'heron') are likely to be attracted to (their nearest neighbour in) the phonological neighbourhood that corresponds to dialect feature A. In this way, feature A (with conditioning environment) is overgeneralized to words that should have feature B (without conditioning environment). This explains why dialect features with a conditioning environment are overgeneralized more frequently than features without a conditioning environment.

We have argued (see section 4.4.4.3) that at first sight, rule-based theory seems to predict a negative effect of conditioning environment on the degree to which dialect features are overgeneralized. The idea is that a conditioning environment offers a clue for the learner (who learns by rule) as to where (i.e. in which environment) a dialect feature applies. The dialect learner will integrate the conditioning environment of a feature as part of the rule that generalizes over this feature. If a child learns where a feature applies, it will be less inclined to overapply the relevant feature, which means that the feature will be overgeneralized less frequently (i.e. negative effect of conditioning environment on overgeneralization). However, as long as the child does not know the structural conditions of a dialect feature exactly, it must learn the relevant feature by trial and error, which entails that the child will make overgeneralization errors (this can be related to the idea of a transitional phase in which overgeneralization is likely to occur). Under this account, rule-based models are also capable of explaining the positive effect of conditioning environment on the degree of overgeneralization.

Since we have shown that the degree of dialect proficiency keeps increasing between the ages of nine and fifteen, we would expect that children gain more insight into the structural conditions of dialect features as they grow older. Since the rule-based prediction is that the acquisition of the exact structural conditions of a particular feature leads to a decrease in the

number of overgeneralizations of that feature, it can be expected (under a rule-based account) that the positive effect of conditioning environment on the degree of overgeneralization weakens as children grow older. In order to find out if this was the case, we tested for the effect of conditioning environment on overgeneralization for the nine-, twelve-, and fifteen-year-olds separately (see tables 7.33A, 7.33B, and 7.33C). Contrary to the (rule-based) expectations, the positive effect of conditioning environment was stronger for the fifteen-year-olds than for the twelve-year-olds, and was stronger for the twelve-year-olds than for the nine-year-olds. This can only be explained from a rule-based perspective if we assume that the fifteen-year-olds do not have a better knowledge of the exact restrictions on the applicability of dialect features than the nine- and twelve-year-olds, which is rather unlikely. A possible exemplar-based account of this result is that neighbourhood effects become stronger as children grow older (i.e. there would be a stronger attraction of new words to 'wrong' phonological neighbourhoods, resulting in overgeneralizations), due to their more developed mental lexicons.

Summarizing, we may conclude that the positive effect of conditioning environment on the degree of overgeneralization can be accounted for by rule-based and exemplar-based models, but the fact that this positive effect becomes stronger as acquisition proceeds, is more difficult to account for in a rule-based model.

8.4.4. Number of competing Standard Dutch variants

We examined the effect of the number of Standard Dutch variants for completeness' sake, but we did not formulate a hypothesis with respect to this effect (see section 4.4.4.4). In the case of competing dialect variants, the relationship with the degree of overgeneralization is more obvious than in the case of competing Standard Dutch variants: it seems plausible that there are more overgeneralizations of one dialect variant at the expense of another variant within large paradigms of competing dialect variants than within small paradigms. An argument in favour of this assumption is the fact that we found 17 different types of overgeneralization within the largest paradigm involved in this study, i.e. the /ɛi/-paradigm. However, it is more difficult to make predictions about the effect of the number of competing SD variants on the degree of overgeneralization.

A significant effect of the factor n SD variants was found. In section 7.3.2.4, we showed that there is a positive effect of the number of competing Standard Dutch variants on the degree to which dialect features are overgeneralized. This positive effect was only significant in the case of the second dialect learners. The positive effect implies that dialect features involving a large(r) number of competing Standard Dutch variants are overgeneralized more frequently than features involving a small(er) number of competing SD variants. Consider the following example. The Maldegem dialect phoneme /ø/ may correspond to five different SD

phonemes: to SD /ɛi/ (e.g. *wijn* 'wine'), to SD /øː/ (e.g. *reus* 'giant'), to SD /œy/ (e.g. *uit* 'out'), to SD /oː/ (e.g. *zoon* 'son'), and to SD /y/ (e.g. *duwen* 'to push'). Such a paradigm of five competing SD variants is the largest paradigm involved in our study. As appeared from table 7.30, two of these dialect features involving five SD variants were particularly frequently overgeneralized, viz. SD /ɛi/ ~ DIA /ø/and SD /œy/ ~ DIA /ø/.[16]

A possible account of this result is that when second dialect learners start to know that, for example, dialect /ø/ matches so many different Standard Dutch phonemes, they lose caution and become overconfident, which may lead to more overgeneralization errors.

8.4.5. Accounting for the effects of geographical distribution

In section 7.3.2.5, it was shown that the factor 'geographical distribution' has a significant negative effect on the degree to which dialect features are overgeneralized. This implies that dialect features with a large(r) geographical distribution are overgeneralized less frequently than features with a small(er) distribution. There was no significant interaction with home language, indicating that the effect of geographical distribution on the degree of overgeneralization is similar for the second dialect learners as for the native dialect speakers.

We account for the negative effect of geographical distribution as follows. Recall from chapter 4 that we related the concept of geographical distribution to the distinction between primary and secondary dialect features (cf. Schirmunski 1930). In Schirmunski's view, primary features are not only considered to be the most local features of a particular dialect (i.e. have the smallest geographical distribution), but also the most salient ones, i.e. salient in the sense of deviating most strongly from the surrounding dialects (and probably from Standard Dutch as well; see section 4.2.7). Probably, this salience is responsible for the high degree of overgeneralization of the dialect features with a small geographical distribution: salient features are likely to 'catch the ear' and they are therefore more readily 'available' to be overgeneralized at the expense of other, less salient features. In the process of dialect acquisition, learners may aim (too strongly) at the primary features of the local dialect because of their salience and in this way overgeneralize these primary features. Consider the example below.

Within the /ɛi/-paradigm the overgeneralization of the feature SD /ɛi/ ~ DIA /ø/ at the expense of the feature SD /ɛi/ ~ DIA /e/ (e.g. in *strijken* 'to iron', *kijken* 'to look', etc.) occurs in 164 out of 640 cases (26%) (among the second dialect learners), whereas the

[16] Overgeneralization of the dialect feature SD /y/ ~ DIA /ø/ did not occur because this feature only occurs in a very limited set of words (which are more or less lexical exceptions): *schuw* 'shy', *duwen* 'to push', *ruw* 'rough', *ruzie* 'quarrel'. Overgeneralization of SD /øː/ ~ DIA /ø/ did not occur either.

reverse (i.e. overgeneralization of SD /ɛi/ ~ DIA /e/ at the expense of SD /ɛi/ ~ DIA /ø/) only occurs in 45 out of 1280 cases (4%) (see table 7.30). The negative effect of geographical distribution on the degree of overgeneralization is reflected in this example: the dialect feature SD /ɛi/ ~ DIA /ø/, which has a very small geographical distribution (= 1; see appendix 4), is overgeneralized more frequently than the dialect feature SD /ɛi/ ~ DIA /e/, which has a geographical distribution of 6 (see appendix 4). We could argue that the feature SD /ɛi/ ~ DIA /ø/ is more salient than the other feature, because, for example, the phonetic distance between /ɛi/ and /ø/ is somewhat larger than between /ɛi/ and /e/ (for /ø/, the feature [+round] must be added). Another indication that the feature SD /ɛi/ ~ DIA /ø/ is a more salient feature of the Maldegem dialect than the feature SD /ɛi/ ~ DIA /e/, is the fact that the former feature is referred to in the shibboleth *vijfenvijftig ijzeren vijzen* 'fifty-five iron screws', whereas there is no shibboleth referring to the feature SD /ɛi/ ~ DIA /e/.

Since we did not examine whether the Maldegem dialect features with the smallest geographical distribution are indeed experienced as the most salient ones, and since we did not examine the relationship between the geographical distribution of features and the phonetic distance between the dialect and the Standard Dutch elements involved, we should be careful in drawing firm conclusions. However, we believe that degree of salience and phonetic distance are factors which may play a role in the results for geographical distribution. Therefore, more research should be done investigating these factors.

8.4.6. Accounting for the effects of token frequency

8.4.6.1. Token frequency at the level of the word

In section 4.4.4.6, we formulated the hypothesis that the infrequent words on our word list would display more overgeneralization errors than the more frequent words on our word list (i.e. a negative effect of token frequency on overgeneralization was expected). Put differently, we expected that the correct dialect variant would more often be repressed by an overgeneralized variant in infrequent words than in frequent words. This hypothesis was based on the literature: it has been observed that analogical processes (e.g. regularization) occur to a larger extent in infrequently used words than in frequently used words (cf. Anttila 1972, 1977; Phillips 1984; Bybee 2002). We also argued that a negative effect of token frequency on the degree of overgeneralization could be best accounted for within an exemplar-based model: frequently used words correspond to frequently activated exemplars.

Furthermore, a dialect variant related to a frequently activated exemplar is likely to be acquired better, so no overgeneralization should occur in the relevant word.

Contrary to our expectations, however, the results with respect to the effect of token frequency on the production of overgeneralizations (see table 7.37) did not reveal a significant (negative) effect. So, the exemplar-based hypothesis was not confirmed by the results.

8.4.6.2. Average token frequency per feature

In section 4.4.4.6.2, we hypothesized that the average token frequency of a dialect feature would have a positive effect on the degree to which the relevant feature was overgeneralized. This hypothesis was in agreement with our hypothesis about the effect of incidence on the degree of overgeneralization: both a high incidence and a high average frequency of usage lead language users to be more frequently exposed to the relevant dialect features. Our assumption was that dialect features to which learners are exposed more frequently – whether because of a high incidence or because of a high average frequency – are overgeneralized more often. So, we expected that features which mainly apply to frequently used words are overgeneralized more often than features which mainly apply to infrequently used words. As discussed in section 8.2.2.2, exemplar-based models as well as rule-based models can account for the effects of average token frequency.

However, contrary to our expectations, section 7.3.2.6.2 (table 7.38 and figure 7.28) revealed that the average token frequency of dialect features had a negative effect on the degree to which features were overgeneralized by second dialect learners and native dialect speakers (there was no significant interaction with home language). This result is difficult to account for. Dialect features which mainly apply to frequently used words are acquired better than dialect features which mainly occur in infrequent words (see section 8.2.2.2). However, a better acquisition of a feature does not necessarily imply that this feature is overgeneralized less frequently than a feature which is acquired less successfully (see our discussion in section 4.4.4.1). So, the more successfull acquisition of features with a high average frequency does not necessarily account for the fact that these features are overgeneralized less frequently. However, another account is not immediately available.

8.4.7. Summary

In this section we have accounted for the effects of feature-related factors on the degree to which dialect features are overgeneralized. The positive effect of incidence on the degree of overgeneralization could be accounted for from a rule-based perspective: language users are more strongly exposed to features with a high incidence than features with a low incidence. The high degree of exposure to particular features leads to the fact that language users are better at forming rules that generalize over these features, and the formation of rules is the

necessary requirement for overgeneralizations to occur. An exemplar-based model predicts that features with a high incidence correspond to large(r) exemplar clouds and these clouds are so strong that they are likely to attract new words, which may lead to overgeneralizations.

The positive effect of conditioning environment on the degree of overgeneralization could be accounted for most straightforwardly within an exemplar-based model: features with a conditioning environment are associated to phonological neighbourhoods and these neighbourhoods can attract new words, sometimes resulting in overgeneralization errors. We have argued that rule-based models may also predict a positive effect of conditioning environment on the degree of overgeneralization at the initial stages of dialect acquisition. These models predict that the positive effect will weaken as acquisition proceeds. However, this prediction was not fulfilled: the positive effect of conditioning environment even became stronger as acquisition proceeded.

The positive effect of the number of Standard Dutch variants on the degree of overgeneralization by the second dialect learners was accounted for by assuming that second dialect learners become overconfident when a dialect variant matches a large number of SD variants.

We have argued that the negative effect of geographical distribution on the degree of overgeneralization is probably due to the effect of salience. The most local features of a dialect are often the most salient ones. We assume that precisely this salience makes these features more prone to overgeneralization.

Finally, we hypothesized that the negative effect of the average token frequency of features on the degree of overgeneralization might be related to the fact that features with a high average frequency are known/acquired better.

The interaction with home language was only significant in the case of the factor incidence, indicating that the other feature-related factors have very similar effects on the degree of overgeneralization of the control group and of the second dialect learners, in spite of the fact that the second dialect learners produce significantly more overgeneralizations (within the /ɛi/-paradigm) than the native dialect speakers. The high degree of similarity between native speakers and dialect learners also appeared from a comparison of stepwise regression analyses: for both groups, incidence appeared to have the largest relative effect on the degree of overgeneralization (see tables 7.39A and 7.39B). However, the effect of incidence became less important when conditioning environment was also added to the model. This indicates that the effect of certain factors may be weakened if there are other factors enhancing or diminishing the degree of overgeneralization. Geographical distribution also had a large relative effect on the degree of overgeneralization in both groups (see tables 7.40A and 7.40B).

Conditioning environment was the only factor that significantly interacted with age. We assume that there might have been more significant interactions with age if we had examined younger children as well.

Recall from section 7.3.3 that age was the only speaker-related factor which had a significant effect on the degree of overgeneralization. It appeared that between the ages of nine and fifteen there was a steady decrease of the number of overgeneralization errors made by the second dialect learners. In the case of the native dialect speakers, however, there was a significantly stronger decrease in the number of overgeneralizations between the ages of nine and twelve than between the ages of twelve and fifteen.

8.5. Summary and conclusion

One of the research questions formulated in chapter 1 was the following: do the extent to which feature-related factors affect the degree of dialect acquisition and the nature of the effects of these factors, reveal whether children acquire a language by rule (as in a rule-based model) or in a word-by-word manner (as in an exemplar-based model)? In this section, we have shown that most of the effects of the feature-related factors (on the degree of dialect knowledge/acquisition and on the degree of overgeneralization) can be accounted for from a rule-based as well as an exemplar-based perspective. Only in one case, was an exemplar-based model clearly preferred over a rule-based one: an exemplar-based model (as opposed to a rule-based model) offers a straightforward account for the positive effect of token frequency on the degree to which individual words were realized with the correct dialect variant. The positive effect of token frequency on the degree to which the dialect variants of individual words are acquired is an argument in favour of the assumption that children (both first and second dialect learners) (at least partly) rely on word-by-word learning. Furthermore, the positive effect of conditioning environment on the degree of overgeneralization could be accounted for by rule-based and exemplar-based models, but the fact that this positive effect appeared to grow stronger as acquisition proceeds, was more difficult to account for in a rule-based model.

However, not all exemplar-based predictions were confirmed. First, it was predicted that phonological neighbourhood effects would be stronger in the case of neighbourhoods which mainly contain frequently used words. This prediction was not borne out in the case of the native dialect speakers: the effect of average token frequency was negative for features with a conditioning environment. Second, the exemplar-based prediction that infrequently used words would show more overgeneralization errors than frequently used words was not confirmed.

Not all predictions made by exemplar-based theory could be tested in the present study. For example, exemplar models predict that the degree to which an exemplar cloud is closely surrounded by other clouds in exemplar space influences the ease or difficulty with which features are learned. However, since our research design was not based on exemplar theory, this prediction could not be tested on the basis of our data. Furthermore, the results that were found for the interaction between conditioning environment and average frequency is only

relevant for the exemplar-based prediction that phonological neighbourhood effects can be stronger in the case of neighbourhoods which mainly contain frequent words, if we assume that the relevant phonological neighbourhoods coincide with the concept of conditioning environment as used in this study.

An important conclusion that should be drawn from our results is that we cannot in all cases decide on the specific influence of a particular factor on the degree of *second* dialect acquisition, for in some cases there was no significant interaction effect with home language. With respect to these factors, we must confine ourselves to general conclusions about the effects on the degree of dialect proficiency of second dialect learners *and* native dialect speakers. We also found that the factors which determine the process of second dialect acquisition, mostly (and often to a stronger degree) affect first dialect acquisition as well. As a result, we can conlude that first and second dialect acquisition are largely guided by the same factors.

Whereas in several cases the factors discussed in this chapter interacted with home language, there were only a few significant interactions with age. Hence, most factors do not have a significantly different effect on one age group as compared to another one. This does not mean that there are no differences in the degree of dialect knowledge or in the degree of overgeneralization between the different age groups. We found that there is a steady increase of the degree of dialect knowledge between the ages of nine and fifteen. There was also a significant decrease in the production of overgeneralizations between the ages of nine and fifteen.

Finally, it should be noted that other factors than the ones involved in this study probably play a role in the process of (second) dialect acquisition as well. The fact that none of the R^2-values of the models presented in chapter 7 were extremely high indicates that we did not enter all factors that determine the degree of dialect knowledge/acquisition.[17] We believe that, in particular, personal factors may be important to explain the remaining variance in our data, such as the child's degree of peer group orientation, the number of friends within the dialect speaking peer group, the degree of contact with relatives who speak the local dialect and the degree to which a child is involved in activities outside school. Furthermore, the feature-related factors salience and phonetic distance, and their possible relationship with the degree of geographical distribution of features, deserve further attention in future research. It would also be interesting to perform more analyses of the interaction effects between different factors (i.e. also between the speaker-related factors, e.g. attitude*gender, etc.), and not only with respect to the degree of dialect knowledge/acquisition, but also with respect to the degree of overgeneralization.

[17] The R^2 values would also have been higher if we had included more interaction effects in our analyses.