

Chapter 9: Conclusions and suggestions for future research

9.1. Conclusions regarding the theoretical background

9.1.1. Conclusions regarding the literature on second dialect acquisition

The most important issues from the literature on second dialect acquisition were discussed in the second chapter of this dissertation. The following main conclusions were drawn on the basis of this literature:

- There are indications that children who move to a new dialect region after a certain age do not acquire the target dialect as successfully as children who move at an early age (cf. Payne 1980; Kerswill 1994, 1996; Chambers 1998). These findings have been related to the idea that there is a critical period in which children are capable of achieving a native-like level of dialect proficiency, and that after this period acquisition becomes much more difficult. In the present study, we selected children from different age groups (i.e. nine, twelve, and fifteen years old), which allowed us to examine whether the dialect proficiency of these children keeps improving until the age of fifteen or not, and whether there is a critical point between the ages of nine and fifteen, after which acquisition slows down.
- An important factor in determining the degree of success in second dialect acquisition is the child's social network, i.e. the degree to which a child is oriented to the peer group and a child's position within the group (which can be more central or more isolated) (cf. Kerswill & Williams 2000; Labov 2001; Berthele 2002). Generally, children with the strongest orientation to their peers, i.e. those who have many friends within the peer group, have higher degrees of dialect proficiency than more isolated children. The use of the local dialect among peers can be a way to react against the norms established by adults (e.g. teachers, parents, etc.). Closely related to the degree of orientation to the peer group is the child's attitude towards dialect (speakers) and its motivation to learn the local dialect (cf. Vousten 1995). The wish to belong to the peer group is a strong motivation to acquire the local dialect. This study has implemented attitude/motivation as one of the independent variables determining the degree of success in (second) dialect acquisition.
- Some dialect features are more difficult to acquire than others. The degree of complexity of dialect features has been related to the degree of predictability (cf. Chambers 1998), i.e. the degree to which the output of a particular dialect feature is predictable. It has been proposed that dialect features/rules with many lexical exceptions and opaque outputs are very complex (cf. the Philadelphia short *a* pattern, see Payne 1980;

Chambers 1998). Further, the degree of complexity of features has been related to the degree to which there is a one-to-one relationship between elements of the first and second language (cf. Vousten 1995; Chambers 1998). However, the literature on second dialect acquisition has not proposed a clear-cut implementation of the concept of complexity. In the present study, different factors which were supposed to contribute to the degree of predictability of dialect features (i.e. incidence, number of dialect variants, number of Standard Dutch variants, conditioning environment, productivity, average token frequency) were all implemented as independent variables. For the implementation of some of these factors (e.g. number of competing dialect/Standard Dutch variants), we elaborated on the ideas described in the literature (e.g. with respect to a one-to-one relationship between L1- and L2-elements; cf. Vousten 1995).

- The literature on second dialect acquisition has not been concerned much with the question whether children rely on ‘rules’ when learning a dialect as a second language, or whether they learn the target dialect in a word-by-word manner. Some suggestions regarding the mental mechanisms underlying second dialect acquisition have been made (cf. Payne 1980; Chambers 1998). In these cases, the (main) assumption is that children make use of rules. The present study focused on the discussion between adherents of rule-based models on the one hand and adherents of exemplar-based models on the other.

9.1.2. Conclusions with respect to acquisition theory

Chapter 3 was devoted to the question whether children rely on rules when learning a second language or learn a second language in a word-by-word manner. Two alternative acquisition theories were discussed, viz. rule-based theory and exemplar-based theory.

- Traditional, rule-based models of language learning assume that children learn the predictable features of a language by forming ‘rules’. Two examples of rule-based models were discussed, which both describe the differences between two varieties of a language. The models of Auer (1993) and Taeldeman (1993) propose a vertical dimension of phonology, in which for each separate language system phonetic surface forms are derived from phonemic underlying representations. This vertical dimension is interpreted from the perspective of Lexical Phonology. Furthermore, the models propose a horizontal dimension of correspondence rules, which relates L1-elements (e.g. segments) and structures to equivalent L2-elements and structures. The basic assumption of these rule-based models is that children rely on these correspondence rules to acquire the phonological features of the target dialect.
- Exemplar-based models reject the idea that children generalize over the predictable features of language by forming rules. Instead, exemplar-based models assume that

language learning proceeds in a word-by-word fashion, i.e. that all words of the target language are stored separately in memory together with idiosyncratic (e.g. phonetic) information. The idea of word-by-word storage allows for a straightforward explanation of frequency effects in language acquisition (cf. Bybee 2001). The arguments that were traditionally quoted against exemplar-based theory (by rule-based theoreticians) have been countered (cf. Gillis et al. 2000). We paid special attention to the argument that exemplar-based models cannot account for overgeneralization errors, and we showed how both rule-based and exemplar-based models can account for overgeneralizations.

The rule-based models of Auer and Taeldeman strongly influenced the design of this study. Therefore, the dialect features involved were described in terms of correspondences between Standard Dutch and dialect elements. However, we did not *a priori* assume the psycholinguistic reality of these correspondences, i.e. we did not commit ourselves to the idea that children actually make use of these correspondences. Since our research design was not based on exemplar-based theory, not all predictions made by an exemplar-based model could be tested in this study. Still, our hypotheses allowed us to test a number of rule-based and exemplar-based predictions, and compare them with each other.

9.2. Conclusions of our study on the acquisition of the Maldegem dialect as a second language

On the basis of a number of hypotheses which were proposed in chapter 4, the present study analysed the effects of a number of feature-related factors on the degree to which different dialect features are known or acquired by the (second) dialect learners. The effects of some of these factors on the degree to which features are overgeneralized were analysed as well. The following main conclusions were drawn from these analyses:

- There is a positive effect of incidence (or type frequency) on the degree to which dialect features are known or acquired (i.e. realized correctly). This implies that features with a high(er) incidence are known/acquired better than features with a low(er) incidence. This positive effect of incidence can be accounted for by a rule-based as well as an exemplar-based model. A rule-based account would start from the observation that children are on average more often confronted with dialect features that occur in a large number of words (i.e. with a high incidence) than with features having a low incidence. As a consequence of a higher degree of exposure to a particular feature, children can more easily form a (correspondence) rule that generalizes over the feature, and in this way the relevant feature is acquired more successfully. An exemplar-based account of the positive effect of incidence runs as follows: features with a high(er) incidence correspond to large exemplar clouds. Since large exemplar clouds are usually

characterized by a high degree of activation, it follows that there is a high degree of activation of the dialect feature that is related to the relevant exemplar cloud, which results in a better acquisition of the dialect feature.

Incidence turned out to have the strongest relative effect (of all feature-related factors) on the degree to which the lexical dialect features are known/acquired by the second dialect learners and by the native dialect speakers.

Incidence also had a positive effect on the degree of overgeneralization, indicating that dialect features with a high(er) incidence are overgeneralized more frequently than features with a low(er) incidence. A rule-based explanation of this is that a higher degree of exposure to features with a high incidence facilitates the formation of a (correspondence) rule. Subsequently, this rule can be overapplied, resulting in overgeneralizations (i.e. according to rule-based theory, rules are a prerequisite for overgeneralization to occur). An exemplar-based account is that exemplar clouds with a high degree of activation (i.e. clouds related to features with a high incidence) are likely to exert attraction to new words, which may lead to overgeneralizations (i.e. when words are attracted to a wrong nearest neighbour).

- We did not find a significant effect of the number of competing dialect variants on the degree of knowledge/acquisition of dialect features. However, this factor appeared to interact significantly with some of the other factors which are supposed to contribute to the degree of predictability of dialect features (i.e. incidence, number of competing Standard Dutch variants, conditioning environment, productivity, and average token frequency). The effect of this factor on the degree of overgeneralization was not tested, but we found that most overgeneralizations in our data concern dialect features which belong to the largest set/paradigm of competing dialect variants involved (i.e. the / ε i/-paradigm).
- There was a negative effect of the number of competing Standard Dutch variants on the degree of dialect knowledge/acquisition: features involving a small number of competing SD variants are known/acquired better than features involving a large number of competing SD variants. This negative effect can be accounted for from a rule-based or from an exemplar-based perspective. A rule-based model predicts that it is difficult to form a (correspondence) rule when there is no one-to-one relationship between the L1- and the L2-elements, and that rule formation (and thus acquisition) becomes more difficult as the number of competing variants increases. Likewise, an exemplar-based model predicts that acquisition becomes more difficult in the case of a high degree of competition among items that belong to one single exemplar cloud (i.e. a heterogeneous exemplar cloud).

The positive effect of this factor on the degree of overgeneralization can be accounted for as follows: if a dialect learner starts to be aware of the fact that a particular dialect

variant matches a large number of competing Standard Dutch variants, the learner may become overconfident and thus make overgeneralization errors.

- There was a positive effect of conditioning environment on the degree to which dialect features are known/acquired. So, features with a conditioning environment are known/acquired better than features without a conditioning environment. Features without a conditioning environment are by definition lexically determined, i.e. they can in principle occur in any phonological environment, but they cannot occur in any word. A rule-based account of this effect is possible. The formation of (correspondence) rules is rather useless in the case of features which are purely lexically determined: even if a learner forms such a rule, he still has to learn, word by word, whether the rule applies or not (i.e. there is no regularity). However, when there is a conditioning environment to which a dialect feature is restricted (i.e. if it can only occur in that particular phonological environment), it makes more sense to form a rule: the conditioning environment can be integrated into the (correspondence) rule, which makes the rule more specific and thus helps to acquire certain dialect features.

The positive effect of conditioning environment was predicted by an exemplar-based model as well. A dialect feature which is restricted to a conditioning environment corresponds to a particular phonological neighbourhood which is formed on the basis of phonological similarity (i.e. it contains words with the same phonological (conditioning) environment). Strong phonological neighbourhoods are likely to attract new words. In this way, the correct dialect realization (i.e. that of the nearest neighbour) can be directly copied onto new words.

However, this operation (i.e. copying the dialect variant of the nearest neighbour onto new words) can also lead to incorrect forms, more precisely to overgeneralizations, i.e. when words are attracted to the wrong nearest neighbour (in heterogeneous neighbourhoods, not all items refer to the same dialect variant). This is an exemplar-based account of the positive effect of conditioning environment on the degree of overgeneralization. We have argued that this positive effect can be accounted for by a rule-based model as well, but only with respect to the initial stages of dialect acquisition, i.e. when the exact structural conditions of dialect features have not yet been acquired. The rule-based prediction is that the positive effect of conditioning environment on overgeneralization becomes weaker as acquisition proceeds, because the restrictions on dialect features are acquired better. However, we found that the effect becomes even stronger, which is rather difficult to account for from a rule-based perspective. This result could be accounted for from an exemplar-based perspective: phonological neighbourhood(s) (effects) become stronger as the mental lexicon develops.

- There was a positive effect of token frequency on the degree of dialect knowledge or acquisition, which indicates that the dialect variant of frequently used words was more often realized correctly than the dialect variant of infrequently used words. Rule-based

models of language acquisition cannot really account for these frequency effects. On the contrary, exemplar-based models offer a straightforward account: since all words have their own representation in memory, it follows that some words/exemplars (in particular, frequently used ones) are activated more frequently than others. This results in greater lexical strength and explains why the dialect variants of some words are acquired better than others.

We did not find a significant effect of token frequency on the degree of overgeneralization, although it has repeatedly been shown in the literature that infrequent words are more liable to analogical processes. Hence, the exemplar-based prediction that infrequent words are more prone to overgeneralization errors was not confirmed.

- There was a positive effect of average token frequency per feature on the degree to which features were realized correctly. The rule-based as well as the exemplar-based account of this effect are very similar to the respective accounts of the effect of incidence. A high average frequency leads to a high degree of exposure to a particular feature which enhances the ease with which a rule is formed (rule-based), and a high average frequency implies a high level of activation of a particular pattern (exemplar cloud), which leads to a better acquisition of the relevant feature.

The negative effect of average frequency on the degree of overgeneralization could only be accounted for by assuming that the positive effect of this factor on the degree of dialect knowledge interferes with the overgeneralization of features with a high average frequency.

The exemplar-based prediction that the effect of conditioning environment (cf. phonological neighbourhood effects) on acquisition would be stronger in the case of a high average frequency (i.e. neighbourhoods consisting for the largest part of frequently used words) was confirmed in the case of the second dialect learners, but was refuted in the case of the native dialect speakers.

- Productivity had a positive effect on the degree of dialect knowledge/acquisition. This means that the productive (in this case, postlexical) features were known/acquired better than the unproductive (in this case, lexical) features. We have argued that productive dialect features may be acquired more successfully because they apply to an infinite number of words (i.e. they have no values for incidence), which implies that language users are frequently exposed to those features. All productive features involved in this study were exceptionless (postlexical) features. Therefore, we have argued that the property of exceptionlessness may also be responsible for the high degree of knowledge/acquisition of the productive features. Rule-based theory predicts that acquisition will be easy when there are no lexical exceptions (i.e. in the case of an automatic rule): one just has to learn the rule and its structural conditions (e.g. the conditioning environment), and one can then apply it to all words that meet these

conditions. In exemplar-based theory, exceptionless features are associated with very homogeneous clouds (i.e. clouds in which all exemplars of type X refer to the same dialect variant), which is a situation that improves the learnability of features.

Since there were no overgeneralizations involving productive (postlexical) features, this variable was of no relevance to the degree of overgeneralization.

- There was a negative effect of geographical distribution on the degree to which features were realized correctly. We argued that this effect is probably related to the property of salience: the features with the smallest geographical distribution (i.e. primary features) are generally the most salient features of a dialect, in that they deviate most strongly from features of surrounding dialects or of the standard language. There was a significant interaction of geographical distribution with age, in that its negative effect was slightly weaker for the oldest age group (i.e. the fifteen-year-olds).

Salience was also considered to be responsible for the negative effect of geographical distribution on the degree of overgeneralization: the most salient dialect features (i.e. those with the smallest geographical distribution) are most readily available to be overgeneralized.

- We found that there were many significant interaction effects between the factors which are supposed to determine the predictability of dialect features (i.e. incidence, number of competing dialect/Standard Dutch variants, conditioning environment, productivity, and average token frequency). These interaction effects 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.

So, most effects of the feature-related factors on the degree of dialect knowledge/acquisition and on the degree of overgeneralization could be accounted for by rule-based as well as exemplar-based theory. Only in two cases was an exemplar-based model preferred over a rule-based one. First, the fact that the positive effect of conditioning environment becomes stronger as acquisition proceeds, is contrary to the prediction made by a rule-based model. Second, the fact that token frequency has a positive effect on the degree to which individual words were realized correctly is an argument in favour of the (exemplar-based) assumption that all words are stored in the mental lexicon together with their phonetic details.

However, not all exemplar-based predictions were confirmed. First, there was no significant effect of token frequency on the degree to which individual words show overgeneralization errors. Second, for the native dialect speakers, the effect of average frequency was negative for features with a conditioning environment, whereas a positive effect was to be expected.

Since we did not start from an exemplar-based design and thus could not test all predictions made by an exemplar model, we should be careful in drawing firm conclusions about which model (rule-based or exemplar-based) explains our findings best. Some of the results could only be accounted for by an exemplar-based model under the assumption that certain conditions were satisfied. For example, the positive effect of conditioning environment can only be explained in terms of phonological neighbourhood effects, when we assume that the responsible phonological neighbourhoods largely coincide with the conditioning environments as defined in this study. However, this is not necessarily the case, since possible phonological neighbourhoods do not necessarily coincide with the conditioning environments assumed in this study. Moreover, the degree of heterogeneity of exemplar clouds can be related to the number of competing (dialect/Standard Dutch) variants as implemented in this study, but it does not completely coincide with it either, since we did not discount the factor ‘conditioning environment’ in our implementation of the factor ‘number of competing dialect variants’.

Another important conclusion is that the factors that guide second dialect acquisition parallel those that guide first dialect acquisition. This conclusion could be drawn from the fact that almost all feature-related factors have similar effects among the native dialect speakers and among the second dialect learners. So, first and second dialect acquisition generally follow the same direction.

Next to the effects of feature-related factors, we also analysed the effects of speaker-related factors on the degree of dialect knowledge/acquisition. The following conclusions were drawn from these analyses:

- We found that the degree of dialect proficiency continues to increase between the ages of nine and fifteen, for the native dialect speakers and the second dialect learners alike. Although it might have been expected that the native dialect speakers reach their ceiling sooner than the dialect learners, this was not the case. We have argued that this may be due to the fact that even the native speakers have no perfect knowledge of the Maldegem dialect (yet). Our data did not reveal a critical turning point in the acquisition process between the ages of nine and fifteen. There was also a significant effect of age on the degree of overgeneralization, in the sense that older children make significantly fewer overgeneralization errors than younger children. For the native dialect speakers, the negative effect of age on the degree of overgeneralization was much stronger between the ages of nine and twelve than between the ages of twelve and fifteen. This was not the case for the second dialect learners.
- There appeared to be a significant effect of gender on the degree of dialect acquisition among the second dialect learners, in that the boys are more successful learners than the girls. This result contradicted Vousten (1995), who did not find significant differences between the dialect learning boys and girls. A possible account for our findings is that

girls attach more value to the higher (overt) prestige of the standard language, and that boys make more efforts to acquire the local dialect through channels other than the family. There was a significant interaction between gender and age: the positive effect of age on the degree of dialect proficiency was stronger for boys than for girls. Gender did not play a significant role in the process of overgeneralization.

- Whereas Vousten (1995) did not find a significant effect of attitude on the degree of success in dialect acquisition, we did find a positive effect of this factor. So, children with a positive attitude towards dialect speakers and use, and with a strong motivation to learn the local dialect actually have a higher degree of dialect proficiency. We argued that the effect of attitude is closely related to the degree of peer group orientation of a child. There was no significant effect of attitude on the degree of overgeneralization.
- There was an (unexpected) negative effect of the origin of the mother on the degree of dialect knowledge/acquisition. This implies that children with a mother from Maldegem are less successful learners of the Maldegem dialect than children with a mother from elsewhere. We suggested that this effect might be due to the fact that Maldegem mothers have more pronounced negative attitudes towards dialect use than mothers from elsewhere, whereas mothers from elsewhere might consider the fact that their children (try to) speak the local dialect as a form of integration.
- As expected, the effect of the origin of the father was positive, indicating that children with a father from Maldegem are more successful learners than children with a father from elsewhere. We have suggested that fathers are more inclined to switch or keep to the local dialect than mothers, even if they have the intention of raising their children in Standard Dutch. Finally, the fact that the effect of (not) having a father from Maldegem is more decisive than the effect of (not) having a mother from Maldegem, might be (partly) due to the fact that we did not include very young children in the research project.

Again, we had to conclude that the speaker-related factors which are important to second dialect acquisition are more or less the same as the ones that guide first dialect acquisition. Both in the case of the native dialect speakers and in the case of the second dialect learners, age was the most important speaker-related factor.

9.3. Suggestions for future research

This dissertation has answered a number of questions, but leaves a number of other questions for further study. The following suggestions can be made with respect to future research:

- Not all factors which are decisive for the degree of success in (second) dialect acquisition were examined in this study. We believe that a large portion of variation

among learners is caused by individual differences among learners with respect to their social networks. Therefore, it would be interesting for future research to examine the degree to which children are oriented towards the peer group and their own position within the group. This could be done by means of a so-called ‘sociogram’ (cf. Labov 2001; Berthele 2002): asking each child for its preferences within the group. Another suggestion would be to do recordings of the interaction between children and their parents. This would allow for a better judgement regarding the precise nature of the home language. Further, more spontaneous speech should be collected in order to be able to distinguish between children’s dialect competence and their dialect performance, since children may be unwilling to use the local dialect despite their knowledge of it.

- Most of the factors discussed did not significantly interact with age. This means that the effects of the relevant factors do not differ between one age group and another. We argued that we might have found significant interactions with age if we had examined children younger than nine. So, the same research should be carried out for younger children as well. Moreover, it would be interesting to perform a longitudinal study, in which a small(er) group of children is followed during a certain period of time.
- This study has suggested that the effects of geographical distribution are most likely related to the (perceptual) salience of dialect features. However, more research should be done to find out whether the most local dialect features (i.e. the primary features) are indeed (experienced as) the most salient ones. Furthermore, the relationship between the geographical distribution of a feature, its degree of salience, and the degree to which it deviates phonetically from equivalent features in surrounding dialects and Standard Dutch, should be investigated profoundly.
- The present study was carried out in Maldegem, which has a transitional dialect between West- and East-Flemish dialects. This transitional status may be a source of (extra) variation. More research should be done to shed light on the question whether the situation is different in other (non-transitional) places.
- Next to the interaction effects (on dialect acquisition) between the factors contributing to the degree of predictability of dialect features, it would also be interesting to examine interactions between the speaker-related factors and to examine interaction effects with respect to the degree of overgeneralization. Further, it might be fruitful to incorporate the factor of conditioning environment in the implementation of the factor of number of competing dialect variants.