John Benjamins Publishing Company

This is a contribution from *Acquiring Sociolinguistic Variation*.
Edited by Gunther De Vogelaer and Matthias Katerbow.

CHAPTER 10

Acquisition of phonological variables of a Flemish dialect by children raised in Standard Dutch

Some considerations on the learning mechanisms

Kathy Rys,* Emmanuel Keuleers,** Walter Daelemans* and Steven Gillis*
* University of Antwerp / ** Tilburg University

This study investigates the learning mechanisms underlying the acquisition of a dialect as a second language. We focus on the acquisition of phonological features of a Flemish dialect by children with Standard Dutch or a regional variety of Dutch as their first language. Data were gathered by means of picture naming and sentence completion tasks. Inspired by Chambers (1992), who found that the data of second dialect learners displayed S-curve patterns which he interpreted as evidence of rule-based learning, we examine whether similar S-curves can be observed in the learner data of our subjects. Contrary to Chambers, our subjects' data do not display S-curves but bear evidence of word-by-word learning across the board. These data are consistent with analogical memory-based models of language acquisition. In order to further investigate the applicability of memory-based reasoning to our data, we perform a computational classification task in TiMBL (Daelemans & Van den Bosch 2005), in which the dialect forms of Standard Dutch words have to be predicted on the basis of various amounts of training data. Not only do we compare the accuracy scores of the model with the acquisition scores of our subjects, the classification task also gives us insight into which words constitute the nearest neighbours of a given word. On the basis of this output, we investigate the effect of the number of enemy neighbours on the degree to which the subjects realize the correct dialect variants of words and on the degree to which they make overgeneralization errors. The major finding of this paper is that dialect forms are more often realized incorrectly and that more overgeneralization errors occur in words with a large(r) number of enemy neighbours.

Keywords: second dialect acquisition, memory-based learning, rule-based learning, neighbourhood effects, overgeneralization

DOI 10.1075/silv.20.10rys

1. Introduction[1]

In recent years, variationist studies have given evidence of an increasing interest in research into second dialect acquisition (cf. Berthele 2002; Chambers 1992; De Vogelaer 2010; De Vogelaer & Rooze-Stouthamer 2006; Katerbow 2012; Kerswill 1994; Payne 1976, 1980; Rys 2003, 2007; Rys & Bonte 2006; Rys & De Valck 2010; Siegel 2010; Tagliamonte & Molfenter 2007; Vousten 1995). In this paper, we use the term 'dialect' as defined by Hinskens (1998: 156): "a linguistic variety, displaying structural peculiarities (often referred to as dialect features) in more than one component, usually of relatively little prestige, lacking codification and mainly used orally in a geographically limited area".[2] The notion of "second dialect acquisition" refers to the acquisition of a dialect as a second language. The qualification "second" indicates that chronologically the acquisition takes place later than the acquisition of the first language (i.e., consecutive as opposed to simultaneous or bilingual language acquisition). Some studies describe situations in which children acquire a dialect at a later age because they moved from one dialect speaking area to another (e.g., Chambers 1992; Kerswill 1994; Payne 1980; Tagliamonte & Molfenter 2007). The situation at hand in our study, however, is somewhat different: the children have always been living in a particular dialect speaking area but they have been raised in a non-dialectal (standard) variety. Hence, they acquire the local dialect at a later age through contacts outside the home situation. Thus, we are dealing with a situation in which there is linguistic variation within one and the same community and in which the different varieties spoken are characterized by different degrees of social prestige (see Section 3.1). In this respect, our study is comparable to that of Vousten (1995), who also examines dialect acquisition in a situation with "Vertikal differenter Varietäten" (i.e., 'vertically distinct varieties', Katerbow 2012: 74, 80). In Flanders, where the current study was conducted, dialect may still be acquired as a mother tongue (i.e., L1), but given the fact that local dialects are spoken less and less in home situations (cf. Hoppenbrouwers 1990; Taeldeman 1989; 1991), children

1. The authors want to thank two anonymous reviewers as well as the editors of this volume for their valuable comments and suggestions on earlier versions of this paper. The research described in this paper was supported by a grant of the F.W.O.-Vlaanderen (Flemish Foundation for Scientific Research) (G.0320.03) and was carried out at the University of Antwerp and Ghent University.

2. In some studies, especially Anglo-American ones, the term of "dialect" is used in a broader sense, that is, to indicate different varieties of a language, such as Canadian English versus British English (cf. Chambers 1992): these different Englishes are equally prestigious and they are both spoken in a geographically large area. In that way, these dialects deviate from our narrower definition of dialect.

increasingly acquire dialect outside the family, e.g., through interactions with their grandparents, other dialect speaking family members, or with peers who do speak the local dialect at home. Peers at school or in youth associations and sports clubs have a large influence on the language of children and youngsters. In spite of the fact that parents may have a negative attitude towards their children speaking the local dialect, the children themselves do not necessarily think of the local dialect as inferior to the standard language. Among the young the prestige of a language variety is much more affected by the general attitude towards that variety within the peer group than by their parents' attitude. Studies in Flanders and The Netherlands (Deprez & De Schutter 1981; Münstermann & Van Hout 1988) have demonstrated that language users generally think of the standard language as posh, formal, detached, etc., while they describe dialects as entertaining, informal, amicable, etc. (cf. "covert" prestige, Labov 2006). This implies that the conditions are favourable for second dialect acquisition. It is not uncommon that children who are raised in the standard language or a regional variety by their parents start learning the local dialect because of their identification with their dialect speaking peers.

In this situation, however, the dialect of the peers has often been affected by processes of dialect levelling, i.e., the process in which local dialects lose their most idiosyncratic features and adopt features from other (surrounding) dialects or the standard language. As a consequence, second dialect acquisition often boils down to so-called "imperfect learning" (cf. De Vink 2004). Given this dialect levelling in the language variety of the peer group, the notion of "second dialect acquisition" should possibly be better replaced by "acquisition of (sociolinguistic) variation" (see Section 3.2), i.e., children acquire variants of the local dialect that are more or less affected by dialect levelling processes. We decided to maintain the term "second dialect acquisition", however, because we want to fall in line with the studies we are starting from and because there are indications that Flemings consider their language situation as consisting of two prototypical language varieties, i.e., dialect and Standard Dutch, irrespective of any intermediate varieties between them (Lybaert 2014, see also Section 3.2).

In this paper, we study the acquisition of phonological features of the dialect of Maldegem (East-Flanders, Belgium) by nine-, twelve- and fifteen-year-olds who were raised in Standard Dutch or a regional variety. More specifically, we focus on the question which learning mechanisms underlie second dialect acquisition: are there indications that children learn by rule, or do they learn the phonological features of a dialect in a word-by-word manner? Research into the learning mechanisms underlying second dialect acquisition is interesting because we are dealing with two language systems (viz., L1 and L2) that are typologically closely related. In the case of our study, we are dealing with a standard language (or a regional variety) as L1 and a dialect, which is roofed over by this standard language, as L2 (thus, vertically distinct varieties; cf. Katerbow 2012: 74). The close relationship between the two systems implies for instance that their lexicons largely overlap. Thus, the L1 and L2 have, for the most part, identical words but those words are pronounced differently. Does this situation promote the formation of rules or does it rather induce word-by-word learning? In spite of the fact that this is an intriguing research question, the literature on second dialect acquisition has barely paid attention to learning mechanisms.

As far as the underlying learning mechanisms are mentioned in the literature on second dialect acquisition, the prevailing view is that an initial stage of lexical or word-based learning is followed by rule acquisition (cf. Chambers 1992). In this paper we investigate whether this view can be maintained for the situation of second dialect acquisition in a Flemish context. In this way we try to plug in to the current discussion between adherents of rule-based theories of language acquisition on the one hand and adherents of exemplar-based models on the other hand.

Rule-based models (also known as abstractionist models, cf. Ernestus & Baayen 2011) assume that language learners make generalizations in the form of abstract representations or rules which are stored separately from the lexicon. After having induced such generalizations from words that are already acquired, language learners can apply these generalizations to new words. Instances of rule-based models are generative phonology (Chomsky & Halle 1968) and, applied to dialectology, the theory of bidialectal phonology (Auer 1990, 1993; Dressler & Wodak 1982; Moosmüller 1988; Taeldeman 1993). According to the premises of bidialectal phonology a learner of the Maldegem dialect will ultimately make a mental rule that stipulates that the Standard Dutch diphthong /ɛi/ corresponds with the dialect monophtong /e/ before a velar or laryngeal consonant (e.g., in *rijk* 'rich', *zwijgen* 'to be silent', etc.). According to exemplar-based models word forms are stored in the mental lexicon without the abstraction of rules. The pronunciation of new word forms is based on analogy with word forms already stored. For example, words with dialect /e/ before a velar or laryngeal consonant are acquired on the basis of their similarity to words that are already stored in the mental lexicon. Instances of exemplar-based models are memory-based language processing (Daelemans & Van den Bosch 2005), analogical modeling (Skousen, Lonsdale & Parkinson 2002) and usage-based learning (Bybee 1995, 2001).

In this paper we examine whether the score graphs representing the acquisition of each of the dialect features concerned display S-curve patterns, because such patterns are interpreted as indications of rule-based learning in an authoritative study on second dialect acquisition (Chambers 1992). Thus, finding such patterns could be an indication that the phonological features of the dialect under consideration are acquired by rule. Furthermore, we examine whether there are indications of lexical learning. By means of a memory-based classification task we map out the most similar lexical neighbours (viz., the so-called "nearest neighbours") of words. This enables us to examine the effects of lexical neighbours on the degree of acquisition of dialect features.

2. Rule-based versus exemplar-based learning

2.1 S-curve as an indication of rule-based learning

Chambers (1992) is an early and leading study into second dialect acquisition. Chambers investigated the acquisition of a number of lexical and phonological features of Southern England English by six Canadian children who had moved to Southern England. He interviewed each of his subjects twice, with an interval of two years. Apart from an informal conversation about their old and new neighbourhoods, his subjects had to evaluate taped accents, identify objects on picture cards and read word lists. On the basis of his observations he proposes eight generalizations, called "principles", which he intends to postulate as "empirically testable hypotheses about the determinants of dialect acquisition" (Chambers 1992: 677). This paper focuses in particular on Chambers' sixth principle: "Phonological innovations are actuated as pronunciation variants" (Chambers 1992: 693), but some of the other principles will also be included in our discussion, more specifically principle 3: "Simple phonological rules progress faster than complex ones" (Chambers 1992: 682) and principle 7: "Eliminating old rules occurs more rapidly than acquiring new ones" (Chambers 1992: 695).

For the underlying learning mechanisms of second dialect acquisition Chambers' principle 6 implies that the acquisition of all phonological dialect variables starts off as lexical learning, i.e., the word-by-word acquisition of a feature. Chambers argues that the variability that can be observed in the acquisition of phonological features is consistent with the theory of lexical diffusion, which states that in processes of sound change a phoneme is altered in one word or a restricted set of words first and only gradually affects other lexical items (cf. Chambers & Trudgill 1980; Chen & Wang 1975; Wang 1969; Wang & Cheng 1970). This kind of sound change typically proceeds slowly in the beginning, then becomes rule-governed and as a consequence spreads rapidly throughout the lexicon, but slows down again towards its completion, sometimes leaving a small residue of words unaltered. Therefore, lexical diffusive sound change is typically represented graphically by an S-curve. Chambers & Trudgill (1980: 177–178) represent the "ideal" S-curve with a gradual spread throughout the first 20% of the lexicon (i.e., 20% of all morphemes/words that meet the structural conditions of the sound change),[3] followed by a sudden and quick spread throughout the following 60% of available items in the lexicon, and closed by a slow spread throughout the last 20% of available words.[4] Obviously, it is rather unlikely that an actual sound change will develop exactly according to these 20–80% boundaries (cf. Devitt 1989: 35).

In his study into the acquisition of Southern England English (henceforth: SEE) by native speakers of Canadian English (CE), Chambers (1992) discovers this ideal S-curve in the score graphs of a number of phonological features: children score either under 20% or over 80% for a particular variable, but none of the children scores anywhere between 20 and 80%. According to Chambers, this pattern indicates that dialect learners acquire the target pronunciation word by word for the first 20% of the instances meeting the structural conditions of a particular phonological feature. However, after a critical mass of words – more specifically, instances of the phonological feature concerned – has been acquired, a rule is acquired that is immediately applied to the following 60% of relevant words, and hence learners do not score between 20 and 80%. Chambers observes this pattern in the score graphs of three of the investigated phonological variables:[5] absence of T-voicing (i.e., voicing of medial /t/ in CE, e.g., in *putting, hearty*), presence of Vowel Backing (i.e., use of a back vowel [ɑ] in SEE, e.g., in *plaster*, *bath*, *dancing*) and absence of Low Vowel Merger (i.e., the low back lax vowels /ɒ/ en /ɔ:/ in SEE are merged as CE /ɒ/, e.g., in *tot/taught*, *offal/awful*).

On the basis of his observation that his subjects either score less than 20% or more than 80% with respect to these phonological variables, Chambers argues that phonological dialect features are initially acquired word by word, but that when enough words have been learned, children acquire a rule. Thus, the acquisition process is a combination of lexical learning and rule-based learning. However, the lexical learning phase is only a transitional stage resulting in rule formation: once a critical threshold of instances of the relevant phonological feature is learned this results in the formation of a phonological rule. In this paper we investigate whether an S-curve also characterizes the score graphs of the phonological features of the Maldegem dialect by nine-, twelve- and fifteen-year-olds. In this way we want to find out whether there are indications of rule learning. At the same time, our

3. It is often morphemes/words with a high token frequency that are affected first by a sound change.

4. Generally, this last 20% consists of morphemes/words with a very low token frequency. Therefore, these infrequent items may sometimes escape the relevant sound change.

5. Next to these three phonological variables, Chambers (1992: 696) describes only two other variables of Southern England English that were investigated in his study, that is, presence of r-lessness and presence of intrusive /r/.

comparison of the learners' data of three age groups allows us to test Chambers' sixth principle, i.e., is there a development from lexical learning in the earlier stages of acquisition to rule learning in the later stages?

Chambers (1992) studied the dialect acquisition process of only 6 subjects. We report on a large-scale study into the acquisition of the phonology of a dialect by non-native speakers of that dialect in which 40 subjects participated (see Section 3 for a discussion of the methodology). Chambers studied children who moved to a new dialect area. In the present study the subjects were born and raised in the dialect area under investigation, but since neither of their parents spoke the local dialect of Maldegem the children were not raised in the local dialect at home. This means that they only come into contact with that dialect later in life and through other people than their primary caretakers. In Flanders, children usually start going to school between the ages of 2;6 and 3;0. This means that all subjects were confronted with other than parental (linguistic) influences from the age of 2;6. In Chambers' study, the 6 subjects were aged between 7 and 15 years old when moving to England.

In addition to these differences in the number of subjects and the ages of the subjects there is another important difference between our study and that of Chambers. All variants in Chambers' study were standard forms, either in CE or SEE, which implies that "'standardization' is irrelevant as a factor in their acquisition" (Chambers 1992: 677), but our study deals with a standardized language variety as L1 and a non-standard dialect as L2. As we already pointed out, this situation involves differences between the overt and covert prestige of the L1 and the L2. Obviously, these are factors that may influence the outcome of the acquisition process.

Because of these differences between our own study and that of Chambers, we must take into account the possibility that different outcomes may result from differences in the input. Nevertheless, Chambers intended his principles to be testable hypotheses about second dialect acquisition and he calls his subjects "a minute sample of the dialect acquirers in any place at any time" (Chambers 1992: 675). Therefore, the aforementioned differences should not keep us from applying Chambers' principle(s) to the language situation under consideration.

2.2 Neighbourhood-effects as an indication of exemplar-based learning

If we observe S-curves in the learners' data of the acquisition of the phonological features under investigation, this would be an indication of rule-based learning: a sudden acceleration in the acquisition of a feature after having learned a critical mass of words is, according to Chambers, an indication of rule learning. On the other hand, if we observe so-called “neighbourhood effects” (i.e., effects of lexical neighbours), this would be an indication of the fact that second dialect acquisition is an exemplar-based analogical process. If the acquisition of a specific dialect feature in a specific word is influenced by the lexical neighbours of that word (i.e., words in the mental lexicon which display the largest similarity with the relevant word), this constitutes an indication of the fact that a new word is classified on the basis of similarity with words already stored in the lexicon. This implies that the dialect pronunciation of the most similar words in the mental lexicon is copied onto the new word. For instance, if a child acquiring the Maldegem dialect wants to pronounce the word *dijk* ‘dike’, he or she will ‘look up’ words resembling *dijk*, such as *lijk* ‘corpse’, *rijk* ‘rich’, *strijk* ‘ironing’, in his or her mental lexicon, and subsequently copy the dialect pronunciation of these words (viz., with dialect [e] instead of Standard Dutch [ɛi]) onto the word *dijk*. Thus, whereas rule-based models of language acquisition are characterized by the premise that abstract generalizations in the form of rules are made during the learning process and that these rules are stored independently from the words which they are based on, analogy-based models do not adopt such abstractions: a new word is not classified by making use of some abstract rule, but by extrapolating the classification (or “class label”; in our case the dialect pronunciation) of the most similar words stored in memory (i.e., the so-called “nearest neighbours”) to the target word.

Among the nearest neighbours, there may be “friendly neighbours” as well as “enemy neighbours”. Friendly neighbours are words with the same dialect pronunciation as the target word, whereas enemy neighbours are words with another classification. Applied to the acquisition of the phonology of the Maldegem dialect: the Standard Dutch word *eik* [ɛik] ‘oak’ is pronounced as dialect [iək], with a centralizing diphthong instead of a closing diphthong. In a classification task in which we want to predict a word’s vowel, the word *eik* should get the class label [iə]. Possible neighbours of *eik* are, among others, *rijk* ‘rich’, *lijk* ‘corpse’, *kijk* ‘look.[6] All of those neighbours have the dialect vowel [e], though; they are pronounced as [ˈreʔə], [lek] and [kek], respectively. In other words, as far as the dialect pronunciation of the vowel is concerned, these words are enemy neighbours of the word *eik*. When a new word is classified on the basis of similarity to an enemy neighbour, overgeneralization occurs. This implies that in this example overgeneralization occurs

6. These examples are all so-called “rhyme neighbours”: the words exhibit similarity with respect to their rhyme (i.e., nucleus + coda). However, in the classification task performed in this study (see Section 2.3), in which a set of 5 nearest neighbours for each test item was considered, not all neighbours were rhyme neighbours (e.g. Figure 1: *huid* ‘skin’ is not a rhyme neighbour of *huis* ‘house’).

if the dialect pronunciation of the enemy neighbours (viz., the class label [e]) is extrapolated to the word *eik*, resulting in *[ek] instead of the correct form [iək].

Finding effects of (enemy) neighbours on the acquisition of dialect variables and on any possible overgeneralizations would constitute a strong indication of exemplar-based learning in second dialect acquisition. In this paper we investigate such effects in the learners' data of the acquisition of the Maldegem dialect. For that purpose we first map out the lexical neighbours of the words we have administered from our subjects by means of a word list. This is accomplished using a classification task in TiMBL, which is discussed in the next section.

2.3 Memory-based language processing

Analogical learning, as described in the previous section, is typical of memory-based language processing. Daelemans & Van den Bosch (2005) developed a computational implementation of exemplar-based analogical learning, which was extended to natural language processing (NLP), that is, the Tilburg Memory-Based Learner or TiMBL. In this study, we use TiMBL for a classification task in which Standard Dutch word forms are presented to the model as test items and the model has to predict the Maldegem dialect variants.

In machine learning a distinction is made between "eager" learning methods (~ rule learning) on the one hand and "lazy" learning methods (~ analogical learning) on the other hand (cf. Aha 1997; Daelemans & Van den Bosch 2005). Eager learning methods abstract generalizations (e.g., probability distributions) from the examples, then get rid of the examples and use the abstract generalizations to process input. Models that are based on analogical reasoning show "lazy" learning, that is, processing of input is delayed until a query is made to the system, then the input is processed by referring to stored examples, but this does not result in some form of abstracted knowledge.

TiMBL is an application of such lazy learning strategies. Another term for lazy learning is memory-based learning, which emphasizes "the role of the storage of all available data" (Daelemans & Van den Bosch 2005: 22). Memory-based learning as implemented in TiMBL is based on storing instances in memory and determining the most similar instances by means of a so-called similarity metric. As was already described in the previous section, the classification or class label of these nearest neighbours is adopted as the classification of the new item. In this way a classification task in TiMBL directly makes use of instances stored in memory instead of deriving generalizations – for example in the form of rules – from these instances.[7]

7. A model that does make generalizations on the basis of the stored exemplars is, for instance, Minimal Generalization Learning (Albright & Hayes 2003). On the basis of a large set

TiMBL proceeds as follows for the classification task in this study: in the learning phase or training phase Standard Dutch word forms (i.e., the so-called training items) are stored in the system with their correct dialect pronunciations, without any form of abstraction, selection or restructuring. In the test phase training items are used as basis for turning new input (viz., the test items) into output, that is, by attributing the correct dialect pronunciation to the test items. This happens on the basis of similarity. The number of nearest neighbours that is involved in this procedure is rendered as *k*. The value of *k* is established in advance. For the classification task in this study the value of *k* was set to 5,[8] which means that the dialect pronunciation of the five most similar training items was taken into account for determining the dialect pronunciation of a test item. Figure 1 represents, for example, the five training items (viz., *buis* 'tube', *sluis* 'lock', *huid* 'skin', *kruis* 'cross', *pluis* 'fluff') that – as far as the dialect pronunciation of the vowel is concerned – are most similar to the test item *huis* 'house', which in the Maldegem dialect is pronounced as [øs].

```
huis  h œy s  class: ø  predicted: ø  { ø 0.8, œy 0.2 }
# k=1,  1 Neighbor(s) at distance: 0.096
#       b œy s { ø 1.0 }
# k=2,  1 Neighbor(s) at distance: 0.106
#       sl œy s { ø 1.0 }
# k=3,  1 Neighbor(s) at distance: 0.110
#       h œy t { œy 1.0 }
# k=4,  1 Neighbor(s) at distance: 0.116
#       kr œy s { ø 1.0 }
# k=5,  1 Neighbor(s) at distance: 0.125
#       pl œy s { ø 1.0 }
```

Figure 1. Nearest neighbours of *huis* 'house' for the classification of the vowel in *huis* ($k = 5$)

From Figure 1 we can derive the following conclusions: the test item *huis* was correctly classified by TiMBL with regard to the pronunciation of the vowel: TiMBL predicts [ø] (see Figure 1: "predicted: ø"), which matches the Maldegem dialect

of exemplars this model creates stochastic rules that are learned during the training phase and applied during testing. With respect to the classification of English past tenses in an experiment based on nonsense words (i.e., so-called wug testing), Minimal Generalization performs better than a purely analogical model.

8. Following standard practice in Machine Learning, the value of k was determined by cross-validation on the training material (using part of the training material as stand-in test data and trying out different values).

pronunciation (see Figure 1: "class: ø"). There is 80% evidence for this classification, whereas there is 20% evidence for a classification as [œy] (see Figure 1: "{ ø 0.8, œy 0.2 }"). The 80% evidence for the prediction [ø] originates from four nearest neighbours each of which has the dialect vowel [ø], viz., *buis* 'tube' (see Figure 1: "b œy s { ø 1.0}"), *sluis* 'lock' (see Figure 1: "sl œy s { ø 1.0 }"), *kruis* 'cross' (see Figure 1: "kr œy s { ø 1.0 }") and *pluis* 'fluff' (see Figure 1: "pl œy s { ø 1.0 }"). The 20% evidence for the prediction [œy] originates from one nearest neighbour that has the dialect vowel [œy], viz., *huid* 'skin' (see Figure 1: "h œy t { œy 1.0 }"). Thus, the dialect vowel of the majority of the nearest neighbours is copied onto the test word, which in this case results in the correct dialect pronunciation. In the example in Figure 1 the four neighbours delivering the correct classification are so-called friendly neighbours, whereas *huid* 'skin' is an enemy neighbour, viz., a word with another dialect vowel than the test word *huis* 'house'.

The training items in the classification task are Standard Dutch word forms and their dialect equivalents which are represented as syllabified sequences of segments that are aligned to each other (see Table 1). The Standard Dutch word forms have maximally two syllables (see Section 3.4), but through a rule of word-final schwa addition in the Maldegem dialect, the dialect forms may consist of three syllables. In each case, syllables are divided into the properties onset, nucleus and coda. The values of these properties are the specific phonetic segments (i.e., phones). Table 1, for example, represents the Standard Dutch form and dialect form of the word *vogel* 'bird' as they are represented in the training data. The symbol "=" means that the value of a specific property remains blank; SD denotes Standard Dutch, DIA denotes dialect.

Table 1. Representation of the training items in TiMBL: coding of the Standard Dutch form and dialect form of the word *vogel* 'bird'

Properties	**Values**	
	SD vogel	**DIA vogel**
Onset first syllable ($onset_1$)	v	v
Nucleus first syllable ($nucleus_1$)	o	ø
Coda first syllable ($coda_1$)	=	=
Onset second syllable ($onset_2$)	ɣ	h
Nucleus second syllable ($nucleus_2$)	ə	ə
Coda second syllable ($coda_2$)	l	=
Onset third syllable ($onset_3$)	=	l
Nucleus third syllable ($nucleus_3$)	=	ə
Coda third syllable ($coda_3$)	=	=

The classification task in TiMBL consisted in predicting the Maldegem segment for each Standard Dutch segment of words that were not part of the training data. Thus, these transpositions took place at the level of the phonetic segment. Only the transpositions that were relevant to the phonological features under consideration (see Section 3.4) were analysed.

As described above, classification in TiMBL is done on the basis of similarity between the test item and the nearest neighbours, i.e., the most similar word forms in the training data. In our case, the five most similar neighbours of the test item were taken into account. For the determination of these neighbours TiMBL makes use of a so-called "overlap metric": the similarity between a test item and all items stored in memory is computed by making the sum of the number of overlapping values. According to this calculation, the word *vogel* 'bird' (v ø = h ə = l ə =) has, for example, 7 overlapping values with the word *kogel* 'bullet', which is pronounced as [ˈkohələ] in the Maldegem dialect and is represented as (k o = h ə = l ə =) in the training data (viz., only the values of the first two properties differ). The word displaying the largest overlap with the test item is the nearest neighbour of that item. Of all training items the nearest neighbour exhibits the smallest distance to the test item. However, since not all properties are equally relevant to the classification task (e.g., the rhyme, i.e., nucleus + coda, may be more important for the classification task than the combination of onset + nucleus), we make use of Information Gain (IG) Weighting (see Daelemans & Van den Bosch 2005: 29–31). By adding this function to the algorithm that computes the overlap metric, each property (e.g., $onset_1$, $nucleus_1$, $coda_1$, etc.) is evaluated individually, in order to estimate how much information the relevant property contributes to the prediction of the correct dialect pronunciation. On the basis of this overlap metric with IG weighting, the classification task in this study selects the five ($k = 5$) most similar training items for each test item. Subsequently, the model attributes the dialect pronunciation that occurs most frequently among these five nearest neighbours to the test item (see Figure 1).

3. Methodology

In this section we discuss the methodology of our study. We discuss consecutively the research location, the selection of the subjects, the recordings, the word list, phonological features and dependent variables, the training and test items and finally, the way in which neighbourhood effects were measured and the way in which we examined the learners' data on the occurrence of S-curves.

3.1 Research location

Our research was conducted in Maldegem, which is a municipality in the outer northwest of the province of East-Flanders (Belgium). The local dialect of Maldegem can be characterized as a transitional dialect between West- and East-Flemish dialects, with some idiosyncratic phonological features (see Section 3.4). The linguistic situation in Maldegem resembles that of other places in Flanders: different language varieties, which cover the whole spectrum from dialect to Standard Dutch, and which are characterized by different degrees of social prestige, are spoken in Maldegem. Thus, there is a situation of linguistic variation within one community.

The Maldegem dialect is still spoken by many people in everyday communication. With respect to the vitality of the local dialect as opposed to the standard language, Maldegem fits in with West-Flemish places, where dialect generally is extremely well represented (cf. Sabbe 2005), as opposed to other East-Flemish places, where dialect usually has a relatively weak position compared to more standardized varieties (cf. Strijkers 1990). In spite of the fact that the Maldegem dialect is still quite vivid, it has suffered a loss of overt prestige in the course of the last decades. As people have come to appreciate the local dialect less, they are gradually replacing it by a more standard variety in an increasing number of situations, including situations of parents talking with their children. The abandonment of the local dialect as the language spoken with one's children at home can be related to the fact that people fear that raising their children in a non-standard variety might be disadvantageous for their children's reading and writing skills at school.

The standard variety which is spoken in Maldegem is Belgian Dutch or Southern Dutch, which is the southern variety of Standard Dutch.[9] In Maldegem, as in all other Flemish places, Standard Dutch is the norm at school (i.e., the language of education). Next to Standard Dutch, a regional variety called *tussentaal* (lit. 'in-between-language') is spoken in Maldegem as well. The use of *tussentaal* as the home language sometimes results from parents' efforts to raise their children in Standard Dutch while failing to reach the standard because of their own dialect backgrounds. Furthermore, for some people *tussentaal* is the actual target variety, as it does not carry the same "posh" connotations as Standard Dutch. Since *tussentaal* covers the whole continuum between dialect and Standard Dutch and is spoken in a variety of social groups, it displays a lot of social variation (cf. Geeraerts 1999). *Tussentaal* also varies geographically (e.g., the regional variety spoken in East-Flanders differs from the one spoken in West-Flanders; cf. Lybaert 2014) because some of the speakers' dialect features persist along with standard features thus

9. In this paper, the notion of Standard Dutch (short: SD) is used to indicate this southern variant.

producing a regionally coloured variety of *tussentaal*. The degree to which dialect features persist in the regional varieties depends on the linguistic level of the features: whereas morphosyntactic dialect features and widely spread lexical features have intruded the regional varieties in Flanders to a large extent (for examples see Rys 2007: 171–172), phonological dialect features barely persist in the regional varieties of Flemish speakers. The only features that persist are a few phonetic features which can be related to the accent of Flemish speakers, such as the extremely open realization of the short vowels /ɪ/, /ʌ/ and /ɛ/ and the close realization of the diphthongs /ɛi/ and /œy/ in the western Flemish region (Rys & Taeldeman 2007). In our study, however, these phonetic features were not included (see Section 3.4).

3.2 Subjects: Selection and categorization

In this paper we discuss the learner data of 40 children who were born and grew up in Maldegem, but whose parents had migrated from other parts of Flanders to Maldegem and raised their children in Standard Dutch or a regional variety. These children were only a subgroup in a large-scale study in which recordings were made of 164 children living in Maldegem (cf. Rys 2007). Of these 164 children, 128 children were raised in Standard Dutch or a regional variety, the remaining 36 children were native speakers of the local dialect and functioned as a control group. Of the 128 children who only came into contact with the local dialect outside the home situation, 40 children had both a mother and a father from outside Maldegem. This implies that these children did not hear the local dialect at home. Consequently the data from these children are most comparable to Chambers' data, which were gathered from children who had moved to a new dialect area. Therefore, we restrict ourselves to these subjects in this paper.

Although we do not include the data of the control group, i.e., the native speakers of the Maldegem dialect, in this study, it is worth mentioning that even this group did not score 100% for all of the phonological features investigated and that also for these subjects there was still some progression between the ages of nine and fifteen (see Rys 2007: 274). Thus, even children who are still raised in the Maldegem dialect, are, to a certain extent, "imperfect learners" of this dialect, in spite of the fact that they have a considerable lead on the second dialect learners (see Rys 2007: 228–235). This imperfect dialect knowledge in the control group is a logical consequence of the processes of dialect levelling and dialect loss in the parents' language. If parents stop speaking the dialect in its most "traditional" form, it is impossible for their children to learn the original dialect and it is quite plausible that these children will keep making progress in their dialect knowledge after the age of nine. Given that the dialect spoken by the peer group is not a stable variety, but one that is subject to various degrees of dialect levelling (depending on personal characteristics like age, gender, attitude, etc.), second dialect acquisition in the Flemish context actually boils down to the "acquisition of sociolinguistic variation", that is, the dialect of the peer group, which sets an example to the second dialect learners, shows a certain degree of variability itself. However, there are good reasons to believe that Flemings do perceive of their language situation as involving two distinct varieties (i.e., prototypical dialect versus prototypical Standard Dutch) which are rarely used in their "pure" form, whereas they lack a prototypical image of the highly varying intermediate variety called *tussentaal* (Lybaert 2014: 91–99). This supports modelling in terms of separate codes (i.e., L1 speakers of Standard Dutch learning "the dialect" as L2) and thus, legitimizes our use of the term "second dialect acquisition". Despite the fact that the native speakers' knowledge of "traditional" Maldegem dialect forms appears to be unstable, these speakers display a positive attitude towards their dialect and a strong motivation to speak it, as appeared from an attitude and motivation test (see Rys 2007: 149, 332), indicating that there is still a drive to speak the local dialect in the peer group. Thus, children who are not raised in the local dialect still have a group of peers who try to speak the local dialect, which will certainly put pressure on them to try to acquire it.[10]

The selection of the subjects was based on a questionnaire filled in by their parents. This questionnaire was distributed in two primary and two secondary schools in Maldegem among the nine-, twelve-, and fifteen-year-olds. All participants were, according to the parents, raised in Standard Dutch or *tussentaal*. As was already pointed out in Section 2.1, Flemish children generally start going to school at the age of 2;6, which means that all children of a particular age group had been confronted with other than parental influences for more or less the same period of time.

There were two reasons why we did not distinguish between children speaking Standard Dutch at home and those speaking a regional variety. First, both groups of children could be considered as (generally) unacquainted with the Maldegem dialect in their home situations, since their parents came from outside Maldegem. Second, the main differences between Standard Dutch and the regional variety are on the level of morphosyntax (see Section 3.1), whereas there are no noticeable differences between both varieties on the level of phonology, except of some phonetic features typical of the local accent, which were not included among the phonological variables concerned. Since we focus on phonology in our study, this

10. Of course, the degree to which children with non-local parents are motivated to acquire the local dialect and as a consequence master this dialect partly depends on personal factors like attitude, various familial pressures, position within the peer group, etc. The effect of such "speaker-related" factors was investigated by Rys (2007) but falls outside the scope of the present paper.

means that children from both groups (i.e., raised in Standard Dutch vs. raised in regional variety) do not undergo very different L1 influences. Because of the fact that only features which can be characterized as 'accent' persist in the regional varieties of Flemish speakers, combined with the fact that we did not consider accent features in our study, we decided not to take into account the actual place of origin of the mother and father.

With respect to age, children of three age groups (in the school year 2003–2004) were selected, viz., nine-, twelve- and fifteen-year-olds. Children younger than nine did not seem to understand the tasks. Therefore, we chose nine-year-old children as our youngest age group. There are indications in the literature that at this age, children are already largely oriented towards the peer group (cf. Labov 2001; Payne 1980; Tagliamonte & Molfenter 2007). Since we expected that accommodation to the local dialect would increase as children became more oriented towards their peers, we also included twelve- and fifteen-year-olds as participants in our study. Sociolinguistic research has revealed that the influence of peers reaches its peak around the age of 15 or 16 (cf. Hill 1981; Hoppenbrouwers 1990; Kerswill 1994, 1996). An overview of the participants in each age group is presented in Table 2.

Table 2. Number of subjects with both parents of non-Maldegem origin, divided by age and gender

	AGE 9	AGE 12	AGE 15
Number of boys	7	9	4
Number of girls	4	10	6
Total	**11**	**19**	**10**

3.3 Procedure

Each participant was administered a picture naming task, in which he/she was asked to give the dialect variant of the pictured object's name, and a sentence completion task. This method of eliciting data allows us to study the participants' knowledge of dialect variants. In order to investigate how well children can actually converse in the local dialect a sociolinguistic interview is needed. Thus, in addition to the more formal tasks, each recording also consisted of a brief conversation between interviewer and subject about school, hobbies, family and friends, etc. In addition, five recordings were made of 30 minutes of spontaneous conversation between age-mates (see Rys 2007: 156). The analyses discussed in this paper, however, are based on the data from the formal tasks only, because only these data allow for a maximum comparability among subjects.[11]

Responses were recorded with a SONY MZ-N707 portable minidisc recorder, and with a SONY ECM-ZS90 Electret condenser microphone. Subsequently, the recordings were digitized and sampled at 44 kHz, 16-bit stereo. During the recordings, the experimenter (i.e., the first author of this paper) spoke the Maldegem dialect herself – being a native speaker of it – in order to create a more or less informal situation in which the subjects would feel confident to use dialect forms. All recordings took place at school during the school year 2003–2004. Not all background noise could be eliminated, but generally, this noise did not disturb the quality of the recordings. All responses were transcribed phonetically afterwards. When more than one answer was given, only the last response was used in the data processing.

3.4 Word list, phonological features and dependent variables

The picture naming and sentence completion tasks were used to administer a word list of 167 words. The word list contained frequent and less frequent mono- and disyllabic words that are representative of about twenty phonological features of the Maldegem dialect,[12] and consisted of 115 nouns, 32 verbs (infinitives) and 20 adjectives. All words were cognates: the Standard Dutch (short: SD) form and the Maldegem dialect variant were phonological variants of the same lexeme (e.g., *krijt* 'chalk': SD [krɛit] vs. dialect [krøt]). The following phonological features were included in our study:

1. Deletion of /l/ and compensatory lengthening of the preceding vowel,
   a. before a pause; e.g., *bal* 'ball': SD [bɑl] vs. dialect [bɑ:]
   b. before a consonant; e.g., *melk* 'milk': SD [mɛlk] vs. dialect [mæ:k]
2. Deletion of /n/ and compensatory lengthening and nasalization of the preceding vowel before an alveolar fricative; e.g., *spons* 'sponge': SD [spɔns] vs. dialect [ˊspõ:sə]

11. Note that Chambers (1992: 676) also describes data that are based solely on picture-card elicitations and phrase-list readings.

12. The selected dialect features represent the target of acquisition for the second dialect learners. In Rys (2007) it appeared that the native dialect speakers (i.e., the control group) did not respond categorically for all of these features. This implies that the target of acquisition has been affected by processes of dialect levelling. As long as the current dialect variants used by young native speakers of the Maldegem dialect have not been systematically inventoried, however, we are forced to rely on older descriptions of the local dialect (cf. Taeldeman 1976; Versieck 1989) in defining the target of acquisition.

3. Glottalization of /k/ between a stressed and unstressed vowel; e.g., *kijken* 'to look': SD [kɛikən] vs. dialect [ˈkeʔə̰]
4. Deletion of /r/ before an alveolar fricative in a restricted set of monomorphemic words; e.g., *worst* 'sausage': SD [wɔrst] vs. dialect [wost]
5. SD /ɛi/ vs. dialect /e/ before a velar of laryngeal consonant (further on the so-called *rijk*-variable); e.g., SD [rɛik] vs. dialect [ˈreʔə]
6. SD /ɛi/ vs. dialect /ø/ before an anterior consonant (*wijn*-variable); e.g., *wijn* 'wine': SD [wɛin] vs. dialect [wøn]
7. SD /ɛi/ vs. dialect /iə/ (*geit*-variable); e.g., *geit* 'goat': SD [ɣɛit] vs. dialect [ˈɦiətə]
8. SD /ɛi/ vs. dialect /æ̞i/ (*kei*-variable); e.g., *kei* 'boulder': SD [kɛi] vs. dialect [kæ̞i]
9. SD /œy/ vs. dialect /ø/ in all positions except word-final position (*duim*-variable); e.g., *duim* 'thumb': SD [dœym] vs. dialect [døm]
10. SD /o/ vs. dialect /ø/ (*zoon*-variable); e.g., *zoon* 'son': SD [zon] vs. dialect [ˈzønə]
11. SD /o/ vs. dialect /uə/ (*roos*-variable); e.g., *roos* 'rose': SD [ros] vs. dialect [ˈruəzə]
12. SD /ʌ/ vs. dialect /ɛ̠/ (*put*-variable); e.g., *put* 'pit': SD [pʌt] vs. dialect [pɛ̠t]
13. SD /ɔ/ vs. dialect /ɛ̠/ (*pop*-variable); e.g., *pop* 'doll': SD [pɔp] vs. dialect [ˈpɛ̠pə]
14. SD /e/ vs. dialect /iə/ (*been*-variable); e.g., *been* 'leg': SD [ben] vs. dialect [biən]
15. SD /e/ vs. dialect /ɛ̠/ before /r/ (*peer*-variable); e.g., *peer* 'pear': SD [per] vs. dialect [ˈpɛ̠rə]
16. SD /ɑu/ vs. dialect /ɑi/ (*kous*-variable); e.g., *kous* 'stocking': SD [kɑus] vs. dialect [ˈkɑisə]
17. SD /a/ vs. dialect /ɔ̞:/ (*maan*-variable); e.g., *maan* 'moon': SD [ma:n] vs. dialect [ˈmɔ̞:nə]
18. SD /a/ vs. dialect /ɛ̠/ before /r/ + alveolar consonant (*paard*-variable); e.g., *paard* 'horse': SD [part] vs. dialect [pɛ̠rt]
19. SD /ɪ/ vs. dialect /æ̞/ (*rib*-variable); e.g., *rib* 'rib': SD [rɪp] vs. dialect [ˈræ̞bə]

Of the phonological features listed above, features 1, 2, 3, 4, 5, 6, 9, 15 and 18 are phonologically conditioned: these features only apply in a particular phonological context (e.g., "before laryngeal or velar consonant"). The remaining features are lexically determined: on the basis of the phonological context, it is unpredictable whether a specific feature is operative or not. In other words, whether the feature is applied, depends solely on lexical factors and not on phonological ones. A special case is feature 17: strictly speaking it is not a phonologically conditioned feature because it can occur in all environments, also in the environment which is restrictive for feature 18 (e.g., *paard* 'horse': SD [part] vs. DIA [pɛ̠rt], as opposed to *baard* 'beard': SD [bart] vs. DIA [bɔ̞:rt]). However, this feature differs from the other lexically determined features in that it has a large scope (i.e., high type frequency, see Section 4.2).

All of the phonologically conditioned features are, in Chambers' terms, "complex" rules. According to Chambers (1992: 682) simple rules are "automatic processes that admit no exceptions", [w]hereas "complex rules have opaque outputs, that is, they have exceptions or variant forms or (…) they have in their output a new or additional phoneme". As was pointed out not only by Chambers (viz., principle 3, see Section 3.1) but also by Payne (1980), these phonological variables are the hardest ones to acquire in the process of second dialect acquisition.

Two dependent variables are involved in our study. On the one hand, did a subject realize the correct dialect variant of a particular word or not? This dependent variable will be denoted as *dialect realization*: is a particular dialect feature applied correctly or not. On the other hand, a subject may realize another variant instead of the dialect variant. Among other things, this divergent variant may be an overgeneralization. The question whether an overgeneralization was produced or not is the second dependent variable, denoted as *overgeneralization*. We use the notion of overgeneralization to imply (1) the application of a phonological feature (e.g., feature 9: SD /œy/ vs. dialect /ø/) in a word that meets the phonological conditions of that feature (i.e., in the case of feature 9: 'all positions except word-final position'), but which constitutes a lexical exception (e.g. *spuit* 'syringe' → dialect [ˈspæ̝itə] or [ˈspiətə], not *[ˈspøtə]), and (2) the application of a feature (e.g., feature 6: SD /ɛi/ vs. dialect /ø/) in a word (e.g., *rijk* 'rich') that does not meet the phonological conditions of that feature (i.e., in the case of feature 6: "before an anterior consonant").

3.5 Training items and test items

The training items of the TiMBL classification were drawn from a database of Standard Dutch words and their "translation" in the Maldegem dialect. This database was established on the basis of the CELEX lexical database for Dutch (Baayen, Piepenbrock & Gulikers 1995). The CELEX database contains 42.380.000 word tokens and is based on a corpus of written language. From this database we selected the mono- and disyllabic monomorphemic words that were known to 2/3 of the 39 participants in a large-scale lexical decision experiment (Keuleers, Diependaele & Brysbaert 2010). The resulting 3,524 word types were translated into Maldegem dialect, taking into account the Maldegem dialect phonology as described in Taeldeman (1976) and Versieck (1989) and the native dialect knowledge of the first author of this paper. In the case of homographs, only the first word form appearing in the database was preserved. Therefore, the training data were further reduced to 2435 words, which corresponded to 9867 phonetic segments.

The Standard Dutch as well as the dialect words were represented phonetically and divided into syllables. In this way, the Standard Dutch words were aligned with their dialect variants (see Section 2.3). Classification in TiMBL took place on the level of the phonetic segment: for the classification of the word *huis* 'house' [h œy s], for example, the model first predicted the first segment [h], subsequently, it predicted the second segment [œy] and finally, it predicted the third segment [s]. Therefore, the five nearest neighbours of *huis* that are selected in the classification of the first segment differ from the nearest neighbours that are selected in the classification of the second or third segment. Because the majority of the variables investigated is vocalic (except for variables 1–4; see Section 3.4), only the classification of the vowel is relevant to this study.

The TiMBL classification was performed three times: with 1000 phonetic segments, 5000 phonetic segments, and 9867 phonetic segments (or 2435 words)[13] as training data. The information about the number of friendly vs. enemy neighbours of the test words was only gathered on the basis of the last classification, that is, the one using the complete database as training data.[14]

The test items, that is, the words of which TiMBL had to predict the dialect variant, were identical to the words that were administered from the Maldegem subjects (see Section 3.4).

3.6 Measuring neighbourhood effects

In this study we want to find out whether neighbourhood effects occur in second dialect acquisition. More specifically, we investigate the effect of the number of enemy neighbours on the degree of acquisition of dialect features, as well as on the degree of overgeneralization of features. As was illustrated in Figure 1, TiMBL generated the five nearest neighbours of each test item. The nearest neighbours that exhibit the same dialect pronunciation as the dialect variant of the test word are the so-called friendly neighbours; the neighbours that display another dialect pronunciation are enemy neighbours of the test word. For each test word we counted the number of enemy neighbours. Subsequently, we tested the effect of the number of enemy neighbours on the realization of the dialect variant (i.e., the dependent variable *dialect realization*) and on the production of overgeneralizations (i.e., the dependent variable *overgeneralization*) by means of binary logistic regression analyses.

13. 9867 segments was the maximum number of available training data.

14. Obviously, the test words did not belong to the training data.

3.7 Determining S-curves in score graphs

In order to determine whether the acquisition of the Maldegem dialect phonological features displays an S-curve pattern, we created score graphs per dialect feature (see Section 3.4) and per age group (nine-, twelve- and fifteen-year-olds, respectively) which revealed how many subjects obtained a particular score (in percent). Subsequently, it was possible – following the example of Chambers – to examine whether the scores were mainly divided over the low and high ends of the graph (that is, less than 20% or more than 80%, respectively) or whether there were also scores in the middle (that is, between 20 and 80%).

4. Results

In this section we first examine whether we can find S-curve patterns in the score graphs of each phonological feature under consideration (Section 4.1). Subsequently, we discuss the results of the TiMBL classification (Section 4.2) and finally, we discuss the effects of the number of enemy neighbours on the degree of dialect realization (Section 4.3.1) and on the degree of overgeneralization (Section 4.3.2).

4.1 Distribution of scores

Table 3 represents the distribution of the scores for the dependent variable *dialect realization* per phonological feature and per age group. A distinction is made between the number of scores of less than 20%, the number of scores between 20 and 80% and the number of scores of more than 80%. In Table 3 we further indicate whether a feature is phonologically conditioned (indicated as “P” and shaded) or lexically determined (“L”).

Since the possibility of rule acquisition only arises in the case of features that are phonologically conditioned (i.e., the lexically determined features have to be stored in the lexicon anyhow), only the results in the shaded cells are relevant for the question whether there are any indications of rule acquisition. Table 3 shows that there are scores between 20 and 80% for each phonological feature and for each age group.[15] Most features even have a relatively high number of scores between 20 and 80%. Thus, contrary to Chambers, who did not find scores between 20 and 80% for the phonological features he investigated, we do find a considerable amount of scores in the central part of the distribution for all phonologically conditioned

15. The only exception is n-deletion in fifteen-year-olds, where all subjects score more than 80%.

Table 3. Distribution of scores per feature for the dependent variable *dialect realization*

Phonological feature	Phonologically conditioned (P) or lexically determined (L)	Number of scores (N)		
		−20%	20%–80%	+80%
Age: 9 (*N* = 11)				
(1) l-deletion	P	8	3	0
(2) n-deletion	P	6	4	1
(3) k-glottalization	P	4	5	2
(4) r-deletion	P	2	9	0
(5) *rijk*-variable	P	5	4	2
(6) *wijn*-variable	P	4	4	3
(7) *geit*-variable	L	8	3	0
(8) *kei*-variable	L	9	2	0
(9) *duim*-variable	P	3	3	5
(10) *zoon*-variable	L	5	4	2
(11) *roos*-variable	L	7	4	0
(12) *put*-variable	L	9	2	0
(13) *pop*-variable	L	9	2	0
(14) *been*-variable	L	7	3	1
(15) *peer*-variable	P	7	4	0
(16) *kous*-variable	L	9	1	1
(17) *maan*-variable	L	4	4	3
(18) *paard*-variable	P	6	3	2
(19) *rib*-variable	L	6	5	0
Age: 12 (N = 19)				
(1) l-deletion	P	5	9	5
(2) n-deletion	P	2	3	14
(3) k-glottalization	P	2	4	13
(4) r-deletion	P	0	18	1
(5) *rijk*-variable	P	3	11	5
(6) *wijn*-variable	P	1	6	12
(7) *geit*-variable	L	3	16	0
(8) *kei*-variable	L	10	6	3
(9) *duim*-variable	P	0	5	14
(10) *zoon*-variable	L	1	15	3
(11) *roos*-variable	L	2	12	5
(12) *put*-variable	L	14	5	0
(13) *pop*-variable	L	11	8	0
(14) *been*-variable	L	3	2	14
(15) *peer*-variable	P	2	11	6
(16) *kous*-variable	L	7	3	9
(17) *maan*-variable	L	0	7	12
(18) *paard*-variable	P	2	14	3
(19) *rib*-variable	L	11	8	0

Table 3. (*continued*)

Phonological feature	Phonologically conditioned (P) or lexically determined (L)	Number of scores (N)		
		−20%	20%−80%	+80%
Age: 15 (N = 10)				
(1) l-deletion	P	2	5	3
(2) n-deletion	P	0	0	10
(3) k-glottalization	P	0	4	6
(4) r-deletion	P	1	9	0
(5) *rijk*-variable	P	0	2	8
(6) *wijn*-variable	P	1	5	4
(7) *geit*-variable	L	2	8	0
(8) *kei*-variable	L	8	1	1
(9) *duim*-variable	P	0	4	6
(10) *zoon*-variable	L	0	7	3
(11) *roos*-variable	L	2	5	3
(12) *put*-variable	L	3	7	0
(13) *pop*-variable	L	4	6	0
(14) *been*-variable	L	1	3	6
(15) *peer*-variable	P	1	6	3
(16) *kous*-variable	L	1	4	5
(17) *maan*-variable	L	2	3	5
(18) *paard*-variable	P	1	8	1
(19) *rib*-variable	L	1	8	1

features involved. So, we are not able to discern any clear patterns – more specifically S-curves – that point in the direction of rule-based learning. Furthermore, there are no apparent differences between phonologically conditioned (P) and lexically determined (L) features: for both types of features a substantial part of the scores lies between 20 and 80%. In a rule-based model, it would be expected that rule formation – and thus, an S-curve pattern – does occur in the case of phonologically conditioned features, but not in the case of features that have to be acquired word by word. However, such a distinction does not emerge from Table 3.

The observation of lexical learning in our data is in line with Chambers' sixth principle in so far as Chambers argues that the initial stages of dialect acquisition are characterized by lexical learning. Thus, our results uncover a legitimate stage of dialect acquisition. However, no subsequent stage of rule acquisition can be discerned from our data, in contrast to Chambers, who did observe S-curves. In Section 5 we further discuss possible explanations of our results and of the differences between our outcomes and Chambers' findings.

4.2 Accuracy in TiMBL classification

In Table 4 the number of test items that were classified correctly by TiMBL are represented per phonological feature as well as the accompanying accuracy scores. These data are based on three TiMBL classifications with 1000, 5000 and 9867 phonetic segments as training data, respectively. The right column shows the percentages of acquisition of the nine-, twelve- and fifteen-year-old subjects for the relevant phonological features. In the second column, it is indicated whether a feature is phonologically conditioned (P) or lexically determined (L). The third column gives the type frequency of the features, which was calculated on the basis of the number of occurrences of a feature in the database of 3524 Maldegem dialect words, which, in turn, was compiled on the basis of the CELEX database (see Section 3.5).

Table 4. Percentage of correct classifications in the TiMBL classification task and percentage of correct dialect realizations per age group

Phonological feature	P/L	Type-freq.	Number of test items	Percentage of correct classification (= accuracy score) by TiMBL with different numbers of training data			Percentage of correct dialect realization in subjects, per age group		
				1000	5000	9867	Age 9	Age 12	Age 15
(1) l-deletion	P	294	10	0%	50%	70%	11%	53%	55%
(2) n-deletion	P	51	3	0%	33%	67%	29%	83%	97%
(3) k-glottalization	P	78	2	0%	100%	100%	41%	74%	77%
(4) r-deletion	P	16	7	0%	57%	57%	18%	23%	47%
(5) *rijk*-variable	P	18	4	0%	75%	25%	27%	47%	72%
(6) *wijn*-variable	P	79	9	78%	89%	89%	46%	75%	60%
(7) *geit*-variable	L	7	3	0%	0%	0%	7%	40%	40%
(8) *kei*-variable	L	15	5	0%	0%	0%	8%	33%	21%
(9) *duim*-variable	P	91	8	100%	100%	100%	54%	83%	77%
(10) *zoon*-variable	L	16	4	0%	0%	0%	34%	53%	66%
(11) *roos*-variable	L	145	8	100%	88%	100%	19%	65%	50%
(12) *put*-variable	L	34	6	0%	0%	0%	6%	16%	30%
(13) *pop*-variable	L	10	7	0%	0%	0%	14%	24%	26%
(14) *been*-variable	L	94	6	0%	50%	67%	27%	75%	72%
(15) *peer*-variable	P	10	3	0%	0%	0%	16%	63%	75%
(16) *kous*-variable	L	5	5	0%	0%	0%	11%	44%	64%
(17) *maan*-variable	L	294	4	100%	100%	100%	42%	77%	56%
(18) *paard*-variable	P	9	4	0%	0%	0%	30%	61%	58%

Table 4 shows that for certain dialect features, TiMBL has an accuracy score of 100%, even with a relatively small set of training data (i.e., 1000 segments). This means that the model predicts the correct dialect pronunciation of all test words. This is the case for the *duim*-variable, the *roos*-variable and the *maan*-variable. Strikingly, each of these features has a relatively high type frequency: the *duim*-variable occurred in 91 words of the Maldegem CELEX, the *roos*-variable in 145 words and the *maan*-variable in 294 words. The relatively high type frequency implies that a large number of instances of these features occurs in the training data. This high type frequency leads to relatively homogeneous neighbourhoods, clustering the words these features apply to, even in the case of a fairly small training set. When a large number of neighbours has the same dialect pronunciation as the test item, this promotes the correct prediction of the pronunciation of the test item (cf. Figure 1). From a Pearson correlation analysis it appeared that there is indeed a significant correlation between type frequency and the accuracy scores with a training set of 1000, 5000 and 9867 segments, respectively (i.e., $r = 0.509^*$, $r = 0.557^*$ and $r = 0.656^{**}$, respectively).[16]

Table 4 shows that some features are not "acquired" at all by TiMBL, not even when the maximum training set is used. It concerns the *geit*-variable, the *kei*-variable, the *zoon*-variable, the *put*-variable, the *pop*-variable, the *peer*-variable, the *kous*-variable and the *paard*-variable. Six out of eight of these variables are lexically determined. Thus, there seems to be a correlation between the accuracy of the model and the question whether a feature is lexically determined or phonologically conditioned. The lack of a conditioning phonological context in lexically determined features implies that the words to which such features apply are not surrounded by a group of friendly neighbours (i.e., homophonous neighbours that all match with the same dialect pronunciation). The fact is that the words to which such features apply do not share a common phonological makeup, hence they do not cluster in a homogeneous neighbourhood. This makes the prediction of the correct dialect pronunciation more difficult. Furthermore, the eight features that are not acquired by TiMBL all have a relatively low type frequency, which implies that only a few words are available as training items for these particular features. Correct classification is hard when the number of test items in ratio to the number of training items is high: e.g., feature 13 applies to 10 words (i.e., type frequency = 10), 7 of which are selected as test items, leaving only 3 instances of the feature in the training data.

Summarizing, the combination of lexical determination and a low type frequency prevents TiMBL from making correct predictions about certain features.

16. Significance: * means $p < 0.05$, ** means $p < 0.01$.

This is mainly due to the fact that these features do not give rise to neighbourhoods of words which all point in the direction of the same dialect variant. Since our model classifies on the basis of the dialect pronunciation that occurs most frequently within the set of the five nearest neighbours, it is obvious that it classifies incorrectly when the set of nearest neighbours is very heterogeneous.

For a number of features the model clearly performs better as there are more training data (e.g., l-deletion, n-deletion, the *been*-variable). This implies that a minimum number of instances of a particular feature is required for correct predictions. In other words, the development of a larger homogeneous neighbourhood promotes the correct prediction of a particular feature. However, in the case of the *rijk*-variable, TiMBL makes 75% correct predictions on the basis of a training set of 5000 segments, whereas it makes only 25% correct predictions on the basis of the complete training set (i.e., 9867 segments). An explanation for this phenomenon awaits further investigation.

Finally, a comparison of the accuracy scores of TiMBL with the subjects' average scores reveals some striking facts. For all features the accuracy scores increase with age. Generally, most progression is made between the ages of nine and twelve, and in many cases the scores already reach their ceiling at the age of twelve.[17] In each case there is progression in the acquisition of the features as a function of age, which does not hold for the results rendered by TiMBL as a function of the amount of training items. Secondly, in the accuracy scores of TiMBL similar ceiling effects show up only in the case of the *wijn*-variable and k-glottalization.[18] Finally, we also performed correlation analyses to test the relationship between type frequency and the subjects' accuracy scores. Significant correlations showed up for the twelve-year-olds (r = 0.193**) and fifteen-year-olds (r = 0.087**), but not in the case of the nine-year-old subjects (r = 0.041). Contrary to the results for TiMBL, where the correlation between type frequency and accuracy grows stronger with an increasing number of training items, the correlation between type frequency and degree of acquisition in our subjects is stronger for the twelve-year-olds than for the fifteen-year-olds. Roughly speaking, the subjects' scores of acquisition behave somewhat differently from TiMBL's accuracy scores. The difference may be attributed to the fact that the maximum number of training data for the classification task in TiMBL (viz., 2,435 words) may be too small to be representative of the lexicon of nine-,

17. These ceiling effects occur in the case of 11 out of 18 features, i.e., l-deletion, k-glottalization, the *wijn*-variable, the *geit*-variable, the *kei*-variable, the *duim*-variable, the *roos*-variable, the *pop*-variable, the *been*-variable, the *maan*-variable and the *paard*-variable.

18. In the case of the *duim*- and *maan*-variables, the maximum score of 100% is reached immediately, that is, with the smallest set of training data. Therefore, we do not consider these results as ceiling effects.

twelve- and fifteen-year-old children acquiring the Maldegem dialect. In order to be able to make a better comparison between the results of a memory-based classification task in TiMBL and the results of the Maldegem subjects, future research could benefit from using a larger training set, for example by the addition of words from the Spoken Dutch Corpus (CGN), which offers the option to select only those words that occur in a particular region (e.g., only words occurring in the speech of East-Flemish speakers).

4.3 Neighbourhood effects

In this section we discuss the effects of *number of enemy neighbours* on the dependent variables *dialect realization* and *overgeneralization*.

4.3.1 *Effect of number of enemy neighbours on dialect realization*

In order to investigate the effect of the number of enemy neighbours on the realization of the dialect variant of a particular word, we performed a logistic regression analysis with *dialect realization* as binary dependent variable and *number of enemy neighbours* as independent variable (or predictor). The results of this analysis are represented in Table 5.

Table 5. Effect of *number of enemy neighbours* on *dialect realization* in nine-, twelve- and fifteen-year-old dialect learners ($N = 40$)

Predictor	B	S.E.	Exp(B)	Significance
N enemy neighbours	−0.241	0.017	0.786	$p < 0.01$
Constant	0.498	0.058	1.646	$p < 0.01$
−2Loglikelihood	5736.751			
Model chi-square	209.964 (df = 1, $p < 0.01$)			

Table 5 displays a highly significant negative effect of *number of enemy neighbours* on *dialect realization* ($B = -0.241$). This means that the probability of realizing the correct dialect variant of a particular word decreases as this word has more enemy neighbours. The logit[19] decreases with 0.241 for each unit of increase in *number of enemy neighbours*, which in this study has a range from 1 to 5. The negative effect of *number of enemy neighbours* on *dialect realization* was to be expected within the context of a model based on analogical learning, since enemy neighbours are

19. The dependent variable in a logistic regression analysis is called the *logit*: logit = natural logarithm of the *odds*; odds = the ratio of the chance of success ($Y = 1$) to the chance of failure ($Y = 0$).

words that are very similar to the test word, but have another dialect pronunciation. Owing to the fact that there is no unequivocal pattern between a word and its neighbours, in the sense that not all of the neighbours point in the same direction, it gets more difficult for a child learning the dialect to acquire the correct pattern of a particular word. This result legitimizes our exemplar-based approach of the data on dialect acquisition.

The box plot in Figure 2 visualizes the distribution of *number of enemy neighbours* for non-dialect realizations (cf. left box) versus 'correct' dialect realizations (cf. right box). From this figure we can deduce that the median of the variable *number of enemy neighbours* is higher for the realizations deviating from the dialect variant (viz., $x = 4$) than for the dialect realizations (viz., $x = 2$). This means that for the non-dialect realizations, half of the produced forms has 0 to 4 enemy neighbours and the other half has 4 to 5 enemy neighbours. On the other hand, half of the 'correct' dialect realizations has 0 to 2 enemy neighbours and the other half has 2 to 5 enemy neighbours. So, this figure replicates the results of Table 5: there are generally fewer enemy neighbours in the case of the 'correct' dialect realizations.

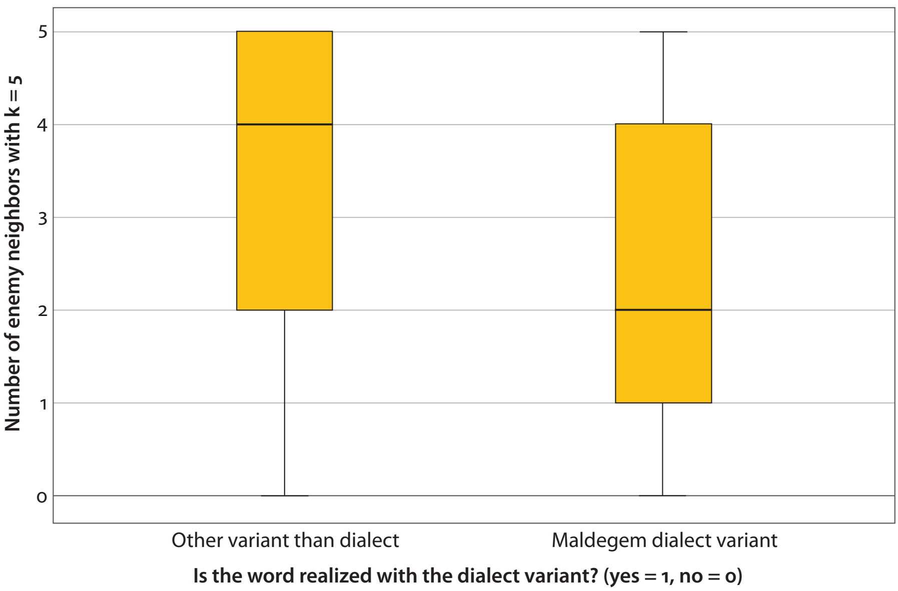

Figure 2. Distribution of number of enemy neighbours of non-dialect vs. dialect realizations

4.3.2 *Effect of number of enemy neighbours on overgeneralization*

In order to investigate the effect of the number of enemy neighbours on the production of an overgeneralization, we performed a logistic regression analysis with *overgeneralization* as binary dependent variable and *number of enemy neighbours* as independent variable. Table 6 gives the results of this analysis.

Table 6. Effect of *number of enemy neighbours* on *overgeneralization* in nine-, twelve- and fifteen-year-old dialect learners ($N = 40$)

Predictor	B	S.E.	Exp(B)	Significance
N enemy neighbors	0.266	0.049	1.304	$p < 0.01$
Constant	−3.698	0.196	0.025	$p < 0.01$
−2Loglikelihood	1648.665			
Model chi-square	32.716 (df = 1, $p < 0.01$)			

Table 6 displays a highly significant positive effect of *number of enemy neighbours* on *overgeneralization*. This means that the probability of producing an overgeneralization in a particular word increases as this word has more enemy neighbours. The logit increases with 0.266 for each unit of increase in *number of enemy neighbours*. This outcome reinforces the results discussed in Section 4.3.1 and was also to be expected within an exemplar-based framework: as there are more words which resemble the test word but which have another dialect pronunciation (i.e., enemy neighbours), the probability of extrapolating an 'incorrect' pronunciation to the test word increases. After all, in exemplar-based models, overgeneralization is explained as the adoption of the class label of an enemy neighbour.

The box plot in Figure 3 displays the distribution of *number of enemy neigbours* for realizations in which no overgeneralization occurs (cf. left box) versus cases in which overgeneralization does occur (cf. right box). The median of *number of enemy neigbours* is lower for those cases in which there is no overgeneralization (viz., $x = 3$) than for the cases in which overgeneralization does occur (viz., $x = 4$). This means that for the cases without overgeneralization, half of the realized forms has 0 to 3 enemy neighbours and the other half has 3 to 5 enemy neighbours. For the cases in which overgeneralization occurs, half of the realized forms has 0 to 4 enemy neighbours and the other half has 4 to 5 enemy neighbours. Thus, in the case of overgeneralization, there are generally more enemy neighbours (cf. Table 6: positive effect of *number of enemy neigbours*).

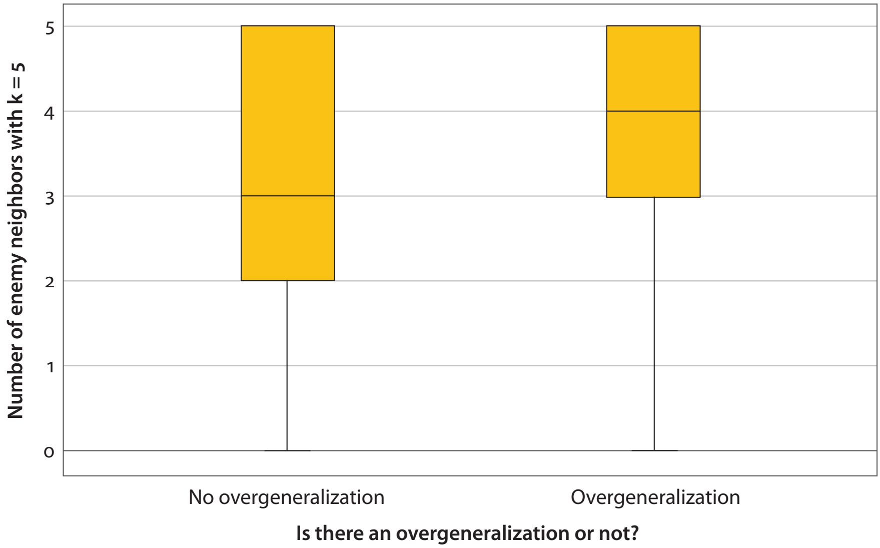

Figure 3. Distribution of number of enemy neighbours for 'no overgeneralization' vs. 'overgeneralization'

5. Discussion and conclusions

5.1 Chambers' sixth principle: Indications of rule learning?

The purpose of this study was to examine whether second dialect acquisition is mainly a matter of rule-based or exemplar-based learning mechanisms. Up until now, the literature on second dialect acquisition has barely paid attention to the question of which learning mechanisms underlie the acquisition of a dialect as a second language. One exception to this is Chambers (1992), who discusses his observations about the acquisition of five phonological processes by six Canadian children who acquire Southern England English as a second dialect. Chambers argues that the phonological features of a dialect are acquired word by word at first, until sufficient words have been acquired to permit rule formation. This idea is based on his observation that S-curve patterns take shape in the score graphs of a number of phonological features. According to Chambers, these S-curves indicate a sudden acceleration in the acquisition of a feature owing to rule formation.

Contrary to Chambers, we did not only include a larger number of subjects (viz., 40), but also a larger number of phonological variables (viz., 19). We have examined whether S-curve patterns become visible in the score graphs of the phonological features of the Maldegem dialect, which was acquired by our subjects as a second language. However, we found that there are scores between 20 and 80% for each feature, which is inconsistent with the idea of an S-curve. Our findings indicate that phonological dialect features are acquired gradually and word by word. That is why a considerable number of dialect learners obtain a score between 20 and 80% for the acquisition of the features considered. Note that this result is not necessarily in conflict with Chambers' principle 6, in which he argues that in the first stages of dialect acquisition phonological features of the new dialect are acquired word by word. To put it another way, lexical learning is a legitimate and predictable stage in second language acquisition. However, the following stage, that is, the stage of rule learning, is missing in our data: for none of the features we observe a sudden rise (i.e., an S-curve) in the score graphs, which would indicate that the subjects had acquired a rule that could be applied to all words falling under its scope. Thus, on the basis of our data, we do not have indications that children learning the Maldegem dialect as a second language rely on rules for phonological dialect features.

With respect to the comparison of our results with those of Chambers, it is important to point at some differences between the designs of both studies, because these differences may be partly responsible for the differences in the outcomes.

Firstly, Chambers did not observe the S-curve pattern for all five phonological features concerned. The only cases in which he found an S-curve pattern were absence of T-voicing, presence of Vowel Backing and absence of Low Vowel Merger. For two other variables, that is, presence of r-lessness (i.e., deletion of non-prevocalic /r/ in words like *summer*, *water*, etc.) and presence of intrusive /r/ (i.e., epenthesis of [ɹ] between vowels at a word or morpheme boundary, like *sofa*[ɹ] *and couch*, *raw*[ɹ] *eggs*, or *draw*[ɹ]*ing*), Chambers did find percentages between 20 and 80%.[20] Remarkably, two of the variables which display S-curves involve the suppression/absence of an L1 feature instead of the acquisition/presence of an L2 feature. This means that the majority of cases in which Chambers observed the S-curve are variables in which speakers had to suppress features of the first dialect. Chambers (1992: 695) himself argues that "eliminating old rules occurs more rapidly than acquiring new ones". Therefore, we should compare our data to those of Chambers with great caution, since the variables involved in our study are all cases of acquiring new dialect features.

With respect to this opposition we tentatively suggest the hypothesis that the elimination of old features and the acquisition of new ones may unfold along different paths of learning. A factor that may affect the learning strategies (i.e., rule-based

20. For presence of r-lessness one out of six subjects scored 30% and for presence of intrusive /r/ one subject scored 40%. Although it concerned only one subject in both of these cases, these data do not seem insignificant to us, given the small number of subjects that were involved in Chambers' study.

or exemplar-based) of second dialect learners, and thus may account for differences between our results and those of Chambers, is how phonemes of the L1 and the L2 relate to each other. For example, in the case of phoneme split, the learner has to acquire a contrast between two phonemes that does not exist in the L1 (so: acquisition of a new L2 phoneme). On the other hand, in the case of phoneme merger the learner has to learn that two different L1 phonemes are represented by one and the same phoneme in the L2 (so: suppression of one L1 phoneme). It is not unthinkable that the nature of the relationship between the L1 and the L2 phonemes is of influence on whether the L2 phoneme is learned on the basis of lexical mechanisms or on the basis of rule formation. Support for this assumption comes from Payne (1980), who observed that the dialect region which children originated from, and thus their first dialect, was of influence on whether these children were more attuned to lexical factors or to rule formation in the acquisition of the Philadelphia dialect short *a* pattern, which can be considered as a complex rule (for an elaborate discussion, see Rys 2007: 35–38). Payne concluded that children who had migrated from the Northern cities to Philadelphia gave evidence of "operating with phonetic rules", whereas immigrants from New York City were "more attuned to lexical factors than rule formation" (Payne 1980: 174). She grounded this conclusion on the finding that the New York City children were more successful in learning the correct Philadelphia realization of the lexical exceptions to the rule of laxing short *a* before /d/ (viz., *mad*, *bad*, *glad*) than in learning the "simple laxing rule" (Payne 1980: 165) in positions before the non-anterior voiceless fricative /š/ (but not before the anterior voiceless fricatives /s, f, θ/). The Northern City children, on the other hand, exhibited the reverse pattern: they were more successful in applying the laxing rule than in acquiring the lexical exceptions.

A second difference between the designs of our own study and that of Chambers is that we study the dialect data of children who grew up in the dialect area under investigation, whereas Chambers focuses on children who moved to the relevant dialect area between the ages of seven and fifteen. Perhaps, this difference may be partly responsible for the differences in outcomes, that is, the presence versus absence of S-curves. One factor that has been discussed in several studies on second dialect acquisition (e.g., Chambers 1992; Kerswill 1994; 1996; Payne 1980) is the age of first contact with the relevant dialect. All of these studies agree on the fact that learning a new dialect is easier before a critical age than after it. However, there is no consensus about the question at what age this critical period of language learning comes to an end. Kerswill (1996) proposes that the critical age of dialect acquisition lies somewhere between the ages of fifteen and sixteen, which is relatively late, but is accounted for as follows: "Adolescents are clearly significant bearers of change; their networks allow them to have wider contacts than younger children, and their desire for a distinct social identity means that they are willing to modify their speech" (Kerswill 1996: 198). If we assume, like Kerswill, that there is a critical age of dialect acquisition somewhere around the age of sixteen, then the fact that some of Chambers' subjects moved to the new dialect area when they were already adolescents, should not be very problematic with respect to the acquisition of a new dialect. However, given the fact that our subjects were exposed to the Maldegem dialect from 2½ years onwards, whereas Chambers' youngest subject was already seven years old when first exposed to the new dialect, there may be some differences in outcomes that can be attributed to the difference in age of first contact. Could it, for example, be the case that younger children are more attuned to lexical factors than older ones? This is, of course, a very tentative hypothesis that should be subjected to further research.

We concluded that the score graphs of the Maldegem dialect features (cf. Table 3) indicate lexical learning instead of rule learning. It could be objected that for some of the phonologically conditioned features (i.e., the only features that allow for rule learning) in our study the number of words showing those features is very low (e.g., type $\text{frequency}_{\text{rijk-variable}}$ = 18, type $\text{frequency}_{\textit{peer}\text{-variable}}$ = 10, type $\text{frequency}_{\textit{paard}\text{-variable}}$ = 9; see Table 4). The fact that these features do not occur in a large number of words makes it less plausible that acquiring 20% of the available instances would be sufficient to acquire the relevant rule. Chambers does not necessarily claim that learning 20% of the words representing a particular rule would be enough to acquire that rule if the type frequency of that rule were very low.[21] However, from Table 3 it appeared that even dialect features with a high type frequency, such as l-deletion or k-glottalization, display a considerable number of scores between 20 and 80%, indicating that these features are learned word by word. In terms of Chambers' assumptions, learning 20% of the available words for these features should be enough to acquire the relevant rule. Therefore, it seems unlikely that the low type frequencies of some of the features investigated account for the absence of S-curves in our data.

In sum, there are a number of differences in the methodology and the data analysed by Chambers and our own study which may obscure the comparison between the results of both studies. However, as far as our data allowed us to test Chambers' sixth principle, we can conclude that our findings are not in conflict with Chambers' suggestion that the initial stages of dialect acquisition are characterized by lexical learning, but that unlike Chambers, we do not observe a sudden acceleration in the acquisition of phonological features which would indicate rule learning.

21. Chambers does not provide information on the type frequencies of his variables, but generally one could say that T-voicing occurs in a large number of words, whereas the more restrictive phonological conditioning of vowel backing and low vowel merger entails a lower type frequency for these variables.

5.2 TiMBL classification and effects of enemy neighbours

In order to further examine the hypothesis that standard speaking children acquire the dialect based on lexical learning strategies, we performed a memory-based classification task in TiMBL, in which the model had to predict the correct dialect forms on the basis of Standard Dutch word forms. This classification task did not only yield information about the accuracy with which an exemplar-based analogical model can predict the right dialect forms of a number of test words, but also information about which are the nearest neighbours of the test words and how many of them are friendly/enemy neighbours.[22] The accuracy scores of a model that is based on a training set of 1000, 5000 and 9867 phonetic segments, respectively, reveal that not only the size of the training set is of importance for a correct classification, but also the presence of a conditioning phonological context (i.e., dichotomy between phonologically conditioned and lexically determined features) as well as the type frequency of the features. Lexically determined features apply to words that are not characterized by an univocal phonological structure, as a result of which they are not clustered in a homogeneous phonological neighbourhood. The words by which they are surrounded, are generally words to which other features apply. In other words, they are surrounded by enemy neighbours. Also type frequency is a factor that plays a role in the formation of neighbourhoods: features that apply to a large number of words will generally be related to a large homogeneous neighbourhood. Given that the classification task in TiMBL is based on the attribution of the most frequent classification from a set of five nearest neighbours to the test item, the model will perform better as the set of nearest neighbours is more homogeneous and consists chiefly of friendly neighbours of the test item.

Not only does neighbourhood appear to play an important part in the prediction of the correct dialect pronunciation by TiMBL, it also turns out to be a crucial concept in the acquisition of phonological dialect features by standard speaking children in Maldegem. We have examined the effects of the number of enemy neighbours on the degree to which standard speaking children learning the local dialect from their peers realize the correct dialect variant and the degree to which they produce overgeneralizations. A larger number of enemy neighbours turns out to have a negative effect on the correct dialect realization of words and a positive effect on the production of overgeneralizations. Both results demonstrate that the acquisition of phonological dialect features is influenced by neighbourhood effects: the more a word is surrounded by words with another dialect pronunciation (i.e., enemy neighbours), the more difficult the acquisition of the dialect pronunciation

22. A purely analogical model like TiMBL is most suited to gain insight into the neighbourhood structure (e.g., how many friendly vs. enemy neighbours) of a set of test items.

of that word will be and the higher the probability of overgeneralization. These results perfectly fit into an analogical exemplar-based model of language acquisition. In such a model, new words are classified on the basis of their similarity with words already stored. As there are more words with another classification (in our case, another dialect pronunciation) among the nearest neighbours, in other words, as a neighbourhood becomes increasingly heterogeneous, the classification of a new item is hampered. Furthermore, the presence of enemy neighbours may account for overgeneralizations: on the basis of similarity to an enemy neighbour, a particular dialect pronunciation is erroneously attributed to a particular word.

These results on the effects of number of enemy neighbours on correct dialect realization and overgeneralization are the strongest indication in our study that lexical or analogical learning is crucial in the acquisition of the Maldegem dialect phonological variables. Moreover, these results legitimize our exemplar-based approach to the data. However, we should note here that our results for the effects of enemy neighbours are interpretable in terms of Chambers' third principle stating that simple rules progress faster than complex ones, especially in the early stages of dialect acquisition (cf. Chambers 1992: 684). Based on his own findings and those of other studies in second dialect acquisition (e.g. Payne 1980; Vousten & Bongaerts 1990; Wells 1973), Chambers argues that complex rules (i.e., rules involving opaque outputs, exceptions and new phonemes) are learned late or never learned at all. Since exceptions to a rule in a rule-based model can be reinterpreted as enemy neighbours in an exemplar-based model, we may say that the difficult acquisition of complex rules in the process of second dialect acquisition is reflected by the negative effect of enemy neighbours on correct dialect realization in terms of an exemplar-based model. Since all the phonologically conditioned dialect features considered in our study are complex, it is perfectly imaginable that the second dialect learners acquire these features at a very late stage or never acquire them perfectly at all. This would explain the negative effects of number of enemy neighbours observed in our data, without excluding the possibility that S-curves, and thus, rule learning would turn up in the acquisition of more simple phonological dialect features.

To conclude, we may say that the acquisition of a dialect as a second language often involves complex language situations in which a variety of factors has to be taken into account. As was suggested by Payne (1980), one of the earliest studies on second dialect acquisition, the acquisition of second dialect phonology may be a combination of attunement to lexical factors and of operating with rules. In this respect, it may be worthwhile for future research to try to apply hybrid models of language processing (e.g., Albright & Hayes 2003; Pierrehumbert 2002), which assume both abstract generalizations and exemplars, to data on second dialect acquisition.

References

Aha, David W. 1997. *Lazy Learning*. Dordrecht: Kluwer Academic Publishers. doi:10.1007/978-94-017-2053-3

Albright, Adam, and Bruce Hayes. 2003. "Rules vs. analogy in English past tenses: A computational/ experimental study." *Cognition* 90: 119–161. doi:10.1016/S0010-0277(03)00146-X

Auer, Peter. 1990. *Phonologie der Alltagssprache. Zur Beschreibung von Variation am Beispiel des Konstanzer Stadtdialekts* [Studia Linguistica Germanica 28]. Berlin & New York: Mouton de Gruyter. doi:10.1515/9783110847574

Auer, Peter. 1993. "Zweidimensionale Modelle für die Analyse von Standard/Dialekt-Variation und ihre Vorläufer in der deutschen Dialektologie." In *Verhandlungen des internationalen Dialektologenkongresses, Bamberg 1990*, ed. by Wolfgang Viereck, 3–22. Stuttgart: Franz Steiner Verlag.

Baayen, Harald, R. Piepenbrock, and Leon Gulikers. 1995. *The CELEX Lexical Database* (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Berthele, Raphael. 2002. "Learning a Second Dialect: A Model of Idiolectal Dissonance." *Multilingua* 21: 327–344. doi:10.1515/mult.2002.014

Bybee, Joan. 1995. "Regular Morphology and the Lexicon. *Language and Cognitive Processes* 10: 425–455. doi:10.1080/01690969508407111

Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511612886

Chambers, Jack K. 1992. "Dialect Acquisition." *Language* 68: 673–705. doi:10.1353/lan.1992.0060

Chambers, Jack K., and Peter Trudgill. 1980. *Dialectology*. Cambridge: Cambridge University Press.

Chen, Matthew Y., and William S.-Y. Wang. 1975. "Sound Change: Actuation and Implementation." *Language* 51: 255–281. doi:10.2307/412854

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.

Daelemans, Walter, and Antal Van den Bosch. 2005. *Memory-Based Language Processing*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511486579

Deprez, Kas, and Georges De Schutter. 1981. "Het Antwerps in Antwerpen. Een attitude-onderzoek." In *Linguistics in Belgium 4. Papers from the Fourth December Meeting of the Belgian Linguistics Circle*, ed. by Marc Dominicy, and Christian Peeters, 17–39. Brussel: Didier Hatier.

De Vink, Leendert. 2004. *Dialect en dialectverandering in Katwijk aan Zee*. Delft: Eburon.

Devitt, Amy J. 1989. *Standardizing Written English: Diffusion in the Case of Scotland, 1520–1659*. Cambridge: Cambridge University Press.

De Vogelaer, Gunther. 2010. "(Not) Acquiring Grammatical Gender in Two Varieties of Dutch." In *Advances in Cognitive Sociolinguistics*, ed. by Dirk Geeraerts, Gitte Kristiansen, and Yves Peirsman, 167–190. Berlin: Mouton de Gruyter. doi:10.1515/9783110226461.167

De Vogelaer, Gunther, and Clasien Rooze-Stouthamer. 2006. "Taalcontact of onvolledige verwerving: casusverlies bij de Zeeuwse pronomina." *Tijdschrift voor Nederlandse Taal- en Letterkunde* 122: 348–368.

Dressler, Wolfgang, and Ruth Wodak. 1982. "Sociophonological Methods in the Study of Sociolinguistic Variation in Viennese German." *Language in society* 11: 339–370. doi:10.1017/S0047404500009350

Ernestus, Mirjam, and Harald Baayen. 2011. "Corpora and Exemplars in Phonology." In *The Handbook of Phonological Theory* [2nd edition], ed. by J. A. Goldsmith, J. Riggle, and A. C. L. Yu, 374–400. Oxford: Wiley-Blackwell. doi:10.1002/9781444343069.ch12
Geeraerts, Dirk. 1999. "Noch Standaard, noch dialect. "Tussentaal" in Vlaanderen en Nederland." *Onze Taal* 68: 232–235.
Hill, J. P. 1981. "De verandering van sociale relaties in de vroege adolescentie." In *Psychologie van jeugdjaren en adolescentie; nieuwe ontwikkelingen en bevindingen*, ed. by Franz J. Mönks, 77–105. Nijmegen: Dekker & Van de Vegt.
Hinskens, Frans. 1998. "Variation studies in dialectology and three types of sound change." In *Linguistics of Variation* [Sociolinguistica 12], ed. by Ulrich Ammon, 155–193. Tübingen: Niemeyer.
Hoppenbrouwers, Cor. 1990. *Het regiolect. Van dialect tot Algemeen Nederlands*. Muiderberg: Dick Coutinho.
Katerbow, Matthias. 2012. *Spracherwerb und Sprachvariation. Eine phonetisch-phonologische Analyse zum regionalen Erstspracherwerb im Moselfränkischen (Wittlich/Eifel)*. Frankfurt am Main: Philipps-Universität Marburg.
Keuleers, Emmanuel, Kevin Diependaele, and Marc Brysbaert. 2010. "Practice Effects in Large-scale Visual Word Recognition Studies: A Lexical Decision Study on 14,000 Dutch Mono- and Disyllabic Words and Nonwords." *Frontiers in Psychology*, doi:10.3389/fpsyg.2010.00174.
Kerswill, Peter. 1994. *Dialects Converging: Rural Speech in Urban Norway*. Oxford: Oxford University Press.
Kerswill, Peter. 1996. "Children and Language Change." *Language Variation and Change* 8: 177–202. Cambridge: Cambridge University Press.
Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.
Labov, William. 2006. *The Social Stratification of English in New York City*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511618208
Lybaert, Chloé. 2014. *Het gesproken Nederlands in Vlaanderen: percepties en attitudes van een spraakmakende generatie*. PhD dissertation, Ghent University.
Moosmüller, Sylvia. 1988. "Sociophonology." In *Variation and Convergence*, ed. by Peter Auer, and Aldo di Luzio, 76–93. Berlin: Walter de Gruyter. doi:10.1515/9783110851601.76
Münstermann, Henk, and Roeland Van Hout. 1988. "Language Attitudes and the Prediction of Dialect Use." In *Language Attitudes in the Dutch Language Area*, ed. by Roeland van Hout, and Uus Knops, 173–188. Dordrecht: Foris. doi:10.1515/9783110857856.173
Payne, Arvilla C. 1976. *The Acquisition of the Phonological System of a Second Dialect*. PhD dissertation, University of Pennsylvania.
Payne, Arvilla C. 1980. "Factors Controlling the Acquisition of the Philadelphia Dialect by Out-of-state Children." In *Locating Language in Time and Space*, ed. by William Labov, 143–178. New York & London: Academic Press.
Pierrehumbert, Janet. 2002. "Word-specific Phonetics." In *Phonology and Phonetics: Papers in Laboratory Phonology* VII, ed. by Carlos Gussenhoven, and Natasha Warner, 101–140. Berlin: Mouton de Gruyter.
Rys, Kathy. 2003. "Secundaire verwerving van fonologische elementen van een dialect." *Taal en tongval* 55: 68–108.
Rys, Kathy. 2007. *Dialect as a Second Language: Linguistic and Non-linguistic Factors in Secondary Dialect Acquisition by Children and Adolescents*. PhD dissertation, Ghent University.

Rys, Kathy, and Dries Bonte. 2006. "The Role of Linguistic Factors in the Process of Second Dialect Acquisition." In *Language Variation – European Perspectives* I, ed. by Frans Hinskens, 201–215. Amsterdam/Philadelphia: John Benjamins. doi:10.1075/silv.1.13rys

Rys, Kathy, and Katrien De Valck. 2010. "Overgeneralisatie in tweede dialectverwerving: het belang van typefrequentie voor de productiviteit van kenmerken en voor het proces van dialectverandering." In *Voor Magda. Artikelen voor Magda Devos bij haar afscheid van de Universiteit Gent*, ed. by Johan De Caluwe, and Jacques Van Keymeulen, 525–540. Ghent: Ghent University, Department of Dutch Linguistics/Academia Press.

Rys, Kathy, and Johan Taeldeman. 2007. "Fonologische ingrediënten van Vlaamse tussentaal." In *Tussen taal, spelling en onderwijs. Essays bij het emeritaat van Frans Daems*, ed. by Dominiek Sandra, Rita Rymenans, Pol Cuvelier, and Peter Van Petegem, 1–9. Ghent: Academia Press.

Sabbe, Saraï Hananja. 2005. *Dialect in Zedelgem: kennis, attitudes, functionaliteit en resistentie*. Unpublished MA thesis, Ghent University.

Siegel, Jeff. 2010. *Second Dialect Acquisition*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511777820

Strijkers, Heidi. 1990. *Taalgebruik, taalattitude en lexicale standaardizering in Melle*. Unpublished MA thesis, Ghent University.

Skousen, Royal, Deryle Lonsdale, and Dilworth B. Parkinson. 2002. *Analogical Modeling. An Exemplar-based Approach to Language*. Amsterdam: John Benjamins. doi:10.1075/hcp.10

Taeldeman, Johan. 1976. *De klankstructuur van het Kleitse dialect*. PhD dissertation, Ghent University.

Taeldeman, Johan. 1989. "De taaltoestand in Vlaanderen." *Neerlandica Wratislaviensia* IV: 77–90.

Taeldeman, Johan. 1991. "Dialect in Vlaanderen." In *Kroesels op de bozzem: het dialectenboek*, ed. by Herman Crompvoets, and Ad Dams, 34–52. Waalre: Stichting Nederlandse Dialecten.

Taeldeman, Johan. 1993. "Dialectresistentie en dialectverlies op fonologisch gebied." In *Dialectverlies en regiolectvorming* [*Taal en tongval* 46], ed. by Frans Hinskens, Cor Hoppenbrouwers, and Johan Taeldeman, 102–119. Amsterdam: Amsterdam University Press.

Tagliamonte, Sali A., and Sonja Molfenter. 2007. "How'd You Get that Accent? Acquiring a Second Dialect of the Same Language." *Language in Society* 36: 649–675. doi:10.1017/S0047404507070911

Versieck, Sabina. 1989. *Het Maldegemse klanksysteem in het heden en honderd jaar geleden*. Unpublished MA thesis, Ghent University.

Vousten, Rob, and Theo Bongaerts. 1990. "Acquiring a Dialect as L2: The Case of the Dialect of Venray in the Dutch Province of Limburg." In *Verhandlungen des internationalen Dialektologenkongresses Bamberg 1990* [Band 4], ed. by Wolfgang Viereck, 299–313. Stuttgart: Franz Steiner.

Vousten, Rob. 1995. *Dialect als tweede taal. Linguïstische en extra-linguïstische aspecten van de verwerving van een Noordlimburgs dialect door standaardtalige jongeren*. PhD dissertation, University Nijmegen.

Wang, William S.-Y. 1969. "Competing Changes as a Cause of Residue." *Language* 45: 9–25. doi:10.2307/411748

Wang, William S.-Y., and Chin-Chuan Cheng. 1970. "Implementation of Phonological Change: the Shuang-Feng Case." *POLA* 2(10): 1–9.

Wells, John C. 1973. *Jamaican Pronunciation in London* [Publications of the Philological Society 25]. Oxford: Blackwell.