Socio-syntax and variation in acquisition

Problematising monolingual and bidialectal acquisition

Leonie Cornips
Meertens Instituut (Royal Netherlands Academy of Arts and Sciences) & Maastricht University

This paper has two aims: first, to emphasize how the linguistic input to which children are exposed is inherently variable and complex. To this end, we will discuss two particular phenomena in Dutch, namely Aux+Inf and gender marking in DP’s. These phenomena lend themselves to a comparison in terms of the nature of the individual, social and regional variation in the input. Second, regarding the question of whether bidialectal acquisition is the same as bilingual acquisition, it seems that there are, in fact, significant differences between the two. Bidialectal children score significantly higher on vocabulary tests than bilingual speakers, they use the Aux+Inf structure in a different way and they also acquire the neuter gender of the Dutch definite determiner significantly faster than bilingual children. Importantly, this paper also explores whether we can maintain a distinction between monolingual and bidialectal children in so-called bidialectal areas.

1. Introduction: Variation and acquisition

When examining the language input which children are exposed to, we can notice tremendous variation in adult and caretaker speech and there is also variation produced by the children themselves. Adult, caretaker and child variation may be the outcome of both internal, linguistic and external, sociolinguistic factors (on the latter, see e.g. Smith et al. 2013 and references cited). Although for many decades examining the role of language input was seen as controversial, nowadays it is

* This work was supported by a Fellowship Grant from The Netherlands Institute for Advanced Study in the Humanities and Social Sciences (NIAS). I am also grateful to NIAS for the correction of my non-native English.
widely accepted that external, input factors simply have to be examined alongside internal, linguistic factors in order to determine the impact of the different types of factors. Several studies have shown that internal and external factors affect distinct domains of language in different ways (Chondrogianni & Marinis 2011, Paradis 2011). There are child internal factors influencing language acquisition rates which vary among individuals, including language aptitude and cognitive maturity as represented by chronological age. For bilingual children, the age of onset (Meisel 2009) and the typological properties of their other language are important (cf. Cornips & Hulk 2006/2008). Child external factors in acquisition research mainly include factors that determine the quantity and quality of the input the child receives in the target language(s). For all children, the quality (i.e. the richness and complexity) of the input can vary depending on the type of so-called native-speaker input they have through activities such as reading (Paradis 2011). For bilingual children, the quantity of the input also varies according to the overall length and intensity of exposure, at home and/or at school (Unsworth 2012; 2013, Unsworth et al. 2011).

1.1 The social configuration of the input

Alongside the quality and quantity of language input, it is also important to consider the make-up of the input from a socio-syntax perspective (e.g. Smith et al. 2013). This approach introduces a number of complexities into the equation. First, sociolinguistic quantitative studies have shown that adult speech is inherently variable in all social uses of language (Labov 1994). Language constitutes not only grammar or an instrument for the communication of messages; it also carries and produces social meanings and social connotations. Numerous urban sociolinguistic studies have shown that the social attributes of speakers correlate closely with the patterns of language variation. Language variation has been revealed not as chaotic but as socially regular. The relevant social attributes of speakers which we can consider to be potential determiners of language variation are, for instance, age, gender, social class, local or ethnic group membership. All of these speaker variables are linked in some way to the individual's place in intersecting social configurations. Linguistic variation is perhaps one particularly salient manifestation of the tendency of groups to construct regional and social identities based on one group's distinctiveness compared to other groups at various levels. While syntax is often viewed within sociolinguistics as a marker of cohesion in large geographical areas, syntactic variants may also act as marker of local identity, as is the case with variability in the phonological component (cf. Cornips 2006). The link between social and linguistic markers therefore has to be understood as very complex and multi-dimensional (Eckert 2000) since language choice and linguistic forms can
index membership to various groups. If categories are seen as flexible and constructed, this may even have an effect on the category of language itself (Schneider, 2014). Seen in this way, the notion of language variation implies that ‘language’ is fluid, heterogeneous and non-discrete and that inter-group, inter-speaker, and intraspeaker variation is a common phenomenon (Benor 2010; Cornips 2006; Eckert 2008). Early sociolinguistic studies, in particular Labov’s (1989) investigation of (-t,-d) deletion in Philadelphia, found that pre-pubescent language learners acquired the socially-situated variability that characterized their parents’ speech patterns. Research by Smith et al. (2007, 2009) on the acquisition of variation in pre-school children with their primary caregivers indicates that sociolinguistic norms are evident from the earliest stages of language acquisition (2013:287). Moreover, Smith et al. (2013) show that caregivers exaggerate the use of standard variants above local ones with younger children but revert to adult-like norms in their speech once the children pass the age of initial acquisition.

It is also necessary to examine language socialization processes if we wish to get an insight into the complex phenomenon of how, and to what extent children acquire their language(s). Children acquire sociolinguistic norms of their communities through participation. That is to say, it is simply by (actively) belonging to their communities, that children acquire how, when and where to use specific linguistic elements. During processes of socialization, children learn to attach specific values to different linguistic elements (Ochs 1993, Ochs & Schieffelin 1995). For example, where a high value is placed on one linguistic element over another, the highly valued element has a better chance of surviving as part of a young child’s individual linguistic repertoire. Crucially, children are particularly tuned into the indexical meanings of grammatical forms which link those forms to the social identities of the interlocutors (Eckert 2008). A telling example is the gender marking in DP’s (see §4 below) i.e. the overuse in Dutch of the common article de ‘the’ instead of neuter het ‘the’ with neuter nouns by bilinguals speaking Berber and Dutch (Cornips et al. 2006; Cornips 2008, Cornips & Hulk 2008). Nortier and Dorleijn (2008:132) describe a interviewee with a Moroccan background in Rotterdam who deliberately uses the ‘wrong’ common definite determiner de ‘the’ with the neuter noun huis ‘home’ instead of the ‘right’ neuter definite determiner het/dat ‘the/that’ because: “in principle you just use the articles deliberately in the wrong way. I am very well aware of the fact that it should actually be het huis [theN houseN/LC], but it would make a dumb impression if I would say dat huis [thatN houseN/LC] out on the street”. This speaker explicitly says he has to deviate from the standard norm to be recognized as someone local, and his overuse of de ‘the’ is indexical of a particular communicative situation (i.e. ‘hanging out’ with friends).

It is, therefore, surely possible to claim that all children are exposed to an input that inherently consists of socially meaningful adult and caregiver intra- and
interspeaker variation. From this perspective, children do not only (have to) acquire a variable target grammar, they also have to acquire (tacit) norms with respect to linguistic variation that are socially meaningful. Thus, during a certain phase, children themselves have to produce a variable outcome that reflects adult and/or peer group and/or local norms. From a socio-syntactical perspective, syntactic variation is always part of a social semiotic process in which a variant functions as a linguistic sign that is indexical of social categories (Eckert 2012).

It is extremely important to further examine the socio-syntactical perspective type within acquisition research, and to investigate whether the child’s use of so-called non-target or ungrammatical forms in the standard language may be a reflection of their social meaningfulness for the child and her community rather than solely a reflection of the child’s lack of grammatical competence or incomplete/unsuccesful acquisition. Indexical sensitivities may account for why and how children may produce non-standard grammatical forms in situated contexts, even if their teachers and/or parents want them to use standard forms. Thus, a variable child language output is not necessarily the outcome of developmental stages per se but it may also reflect and construct social identities in the local communities (cf. Cornips 2008; Kerswill 1996; Smith et al. 2007; 2009; 2013).

1.2 Language standardization, standard language and bidialectal areas

In general, acquisition studies focus on target grammars that belong to a standard or national language. Since codification of a standard language like English, Dutch, and French was (and is) regarded crucial in the formation of nation-states, especially since the end of the nineteenth century, ideas about the standard language play a dominant role in the construction of national identities. A deeply entrenched language ideology is built on the unquestioned principle that a nation-state should be linguistically homogeneous and that a standard language is a discrete, homogeneous linguistic object (Auer 2007, Cornips et al. forthcoming, Kroskrity 2000; Milroy 2001). Thus, historically in language standardisation, linguistic homogeneity has been made salient by downplaying or erasing linguistic variation (Gall & Irvine 1995). A good example is the use of auxiliary gaan ‘go’ in standard Dutch. According to Dutch grammar (see Geerts et al. 1984), the auxiliary gaan ‘go’ may express inchoative aspect ‘is about to’ and/or it may also convey the future reading ‘is going to’, as shown in (1a) and (1b), respectively. The participant in (1) intentionally performs the activity expressed by the predicate.

(Inchoative) Aspect auxiliary:

(1) a. Hij gaat nadenken.
    he   goes   think.INF
    ‘He is about to think about it.’
Tense (future reading rooted in the present) Auxiliary:

b. *Hij gaat verhuizen.*
   he goes move-INF
   'He will move.'

However, Lalleman (1986:63, 74) observes in her spontaneous speech data between caregivers and children that caregivers may also make use of *gaan*+INF to indicate present tense or on-going events, as illustrated in (2b):

(2) a. *Wat gaat het hondje doen?*
   what goes the doggie do?
   'What is the doggie doing?'

(2) b. *Het hondje gaat blaffen!*
   the doggie goes bark
   'The doggie is barking!'

The specific genre of child-directed speech is not present in the codified grammar(s) of Dutch, revealing that these grammars do not exhaustively describe the full repertoire of the *gaan*+INF construction in Dutch. There are many more examples of local, regional or non-standard spoken speech that are not present in codified grammars. One interesting example is the *was/were* alternation as instantiated in Buckie English spoken in north Scotland (cf. Adger & Smith 2005):

(3) a. *He says I thocht you were a diver or somethin*
   'He said 'I thought you were a diver or something.'

b. *Aye, I thocht you was a scuba diver.*
   'Yes, I thought you were a scuba diver.'

According to Poplack and Dion: “It follows that the normative tradition wields little authority over the spoken language. (...) where prescription is diametrically opposed to community patterns (...) the cost of aligning with the standard would be too great for the speaker, who must conform, at least minimally, to the norms of her speech community” (2009:581–582). In acquisition studies, however, it is almost always assumed that adult grammars equal codified grammars and hence national norms. Consequently, it is assumed that children orient to standard forms as the only target forms, as in examples (1) and (3a), whereas local forms that are most often non-standard variants, for example (2) and (3b) tend to be ignored (but see Smith et al. 2013).

This paper tries to show that the target in the acquisition process is, of course, spoken speech, and this is inherently variable since ‘local’ forms exist alongside ‘standard’ (national) forms. Due to the standard language ideology, it is too easy to consider adult speakers in dominant areas in contemporary western communities to be the ‘true’ monolinguals, or native speakers of a language (whatever that
means anyway, in an era of globalization cf. Blommaert 2010). Yet most, if not all competent speakers of a ‘language’ can usually resort to a range of repertoires including features that vary from (what is considered) standard to non-standard depending on what is needed for socially significant interaction, and on social and discourse contexts. The systems of adult speakers are not static but rather are part of on-going processes of social dynamics (Cornips & Corrigan 2005). Indeed, even in those increasingly rare communities in which supralocal models are absent, face-to-face interactions are often polylectal (cf. Auer 2005).

To give some initial context, the focus in this paper is on an area where language variation is even more complicated than in a so-called monolingual standard speaking one, namely a bidialectal area where adults speak both the standard language and the local dialect. The area under question is the province of Limburg located in the south-eastern part of the Netherlands, bordering both Belgium (in the west) and Germany (in the east). Bidialectal children in Limburg grow up in families whose language choice patterns are reflected in the surrounding bidialectal communities. Unlike expatriate families, this is rarely a ‘one parent, one language’ setting but rather language choice patterns are context-dependent, relying e.g. on particular interlocutors within the family domain, topics, and activities (Cornips & Hulk 2006). The speech repertoire in Limburg is of the so-called intermediate type which, according to Auer (2005) is presumably the most widespread in Europe today. In this type of repertoire, there is a very close structural relationship between the standard variety and the dialects and there is no longer a clear-cut separation between the language varieties. Speakers can change their way of speaking without a clear and abrupt point of transition between these varieties and standard forms can be found in local dialects while dialect forms can also be found in the standard variety (Cornips 2006, 2014). Every emerging intermediate speech repertoire is the result of induced language contact situation and/or processes of standard – dialect and/or dialect – dialect convergence or vertical and horizontal levelling, respectively (cf. Cornips & Corrigan 2005). Moreover, it is important to note that, since families and the surrounding communities are bidialectal, the distinction between monolingualism and bilingualism here is blurred (cf. Auer 2007). Each so-called monolingual standard speaker in this area has a passive knowledge of more syntactic alternatives than (s)he actually uses due to the fact that standard, dialect and intermediate variants can be heard every day in his/her community (cf. Cornips 2006). Of course, the same also holds true for children growing up in this type of community. On the other hand, bilingual children in Europe often grow up in families whose language choice patterns are not mirrored in the surrounding community (Cornips & Hulk 2008).

When focusing on such a bidialectal area, the question therefore arises whether bidialectal acquisition is of the same type as bilingual acquisition and whether it
is relevant to make a distinction between bidialectal children and monolingual children (or vice versa). Furthermore, it is to be expected that child bidialectal acquisition is of a different type to child bilingual acquisition. The focus of this paper is to find out whether this expectation is true.

This paper will present acquisition data for monolingual, bilingual and bidialectal children and adults. It is organized as follows. In Section 2, vocabulary scores for monolingual, bilingual and bidialectal children from the same city in Limburg are presented. These vocabulary scores will provide the first evidence as to whether bilingual children acquire Dutch in a similar way as bidialectal children. In Section 3, the Aux+Inf construction will be presented, appearing to show huge variation in child and adult standard Dutch. Section 4 details the acquisition of the neuter gender of the definite determiner in Dutch as this neuter gender is difficult to acquire for both monolingual and bilingual children. In this respect, this phenomenon counts as a diagnostic to see whether bilingual acquisition is of the same type as bidialectal acquisition. Finally, Section 5 contains the discussion of the above issues.

2. Bidialectal, bilingual and ‘in-between’ acquisition of vocabulary

The phenomenon of bidialectalism in The Netherlands has increased since the latter half of the former century, and monolingual speakers of dialects who do not speak the standard language have become the exception. Nowadays, children often acquire Dutch in addition to the local dialect and are therefore raised bilingual or rather bidialectally, either from birth (2L1, early L2) or from school age onwards (child L2). In Limburg there has been a long period of contact between Dutch and the local dialects (about 100 years) and this contact has been intensive. This area is considered a peripheral and non-standard language area from the perspective of its inhabitants as well as from the dominant economical, political and cultural western region of the country, the so-called Randstad area where Amsterdam, Rotterdam and The Hague are located. Nowadays, around 900,000 people or 75% of the inhabitants in Netherlandic Limburg claim to speak a Limburg dialect (Driessen 2006: 103). The Limburgian dialects maintained a far stronger social and cultural position here than other dialects elsewhere in the Netherlands (cf. Cornips 2013a and references cited there). Moreover, speaking a Limburgian dialect is not evaluated as a characteristic of lower social status (Driessen 2006), but as a (positive) expression of regional or local loyalty.

Let us first address the issue of whether bidialectal children behave more like bilingual or monolingual children in terms of vocabulary acquisition. To test this, 49 children between the ages of 4;1 and 6;0 were selected from pupils in two
primary schools in Weert (Limburg) (cf. Dirkx 2012). The children are classified as monolinguals \((n = 13)\), bidialectals \((n = 20)\) and bilingual children \((n = 16)\). The ‘monolinguals’ label refers to children who grow up in a household with only standard Dutch as a home language. Naturally, these children will encounter the local dialect in the surrounding communities so they cannot be considered ‘pure’ monolingual. However, this is of course true for all communities in contemporary western societies where all kinds of languages can be heard on a daily basis. The bidialectal children have all been exposed to a Limburgian dialect from birth onwards whereas the group of bilinguals consists of children with the following language backgrounds: Moroccan Arabic/Berber \((n = 11)\), Turkish \((n = 2)\), Polish \((n = 1)\), Dari \((n = 1)\) and Rumanian \((n = 1)\). Note, however, that some of the Moroccan–Arabic/Berber children have acquired the local dialect as well. Thus, the local dialect is just one of a wider set of resources available for everyone to choose from (cf. Cornips 2013a).

The children in this study were classified according to their scores on the standardised Peabody Picture Vocabulary Test (PPVT- III-NL) (Dunn et al. 2005; Dunn & Dunn 2007). According to Unsworth and Hulk (2010), this test is a standardized vocabulary task and thus can be used as a general indicator of children’s relative proficiency in Dutch. In their study, Unsworth and Hulk classified children in groups: Group I was those with an WBQ below mean (95 or lower); children with a mean WBQ (between 95 and 108) were Group II and children with a WBQ above mean (108 or higher) as Group III. The present study follows the lead of Unsworth and Hulk (2010), in assuming that vocabulary development can be taken as a general indicator of children’s overall linguistic development. It is also assumed that Group III is at a more advanced stage of linguistic development than Group II, and Group II is more advanced than Group I. The scores for the three groups are presented in Table 1, along with the children’s ages.

Table 1 shows that 69 percent of the bilingual children were in Group I, the least advanced group, as opposed to only 15 percent of the bidialectal children. When it comes to the most advanced children, only 6 percent of the bilingual children were in group III, as opposed to 45 percent of the bidialectal children. These two different distributions are a significant token that bidialectals do differ from bilinguals in terms of linguistic development. Moreover, bilinguals reveal lower scores than the monolingual children whereas this is not the case for the bidialectals. Hence, 45% of the bidialectals are classified in Group III compared to 31% of the monolinguals. It seems that bidialectals perform slightly better than monolinguals in the Peabody Picture Vocabulary Test. However, this is certainly not the whole story. Regarding the sociolinguistic context in which the children grow up, both monolingual and bidialectal (and bilingual) children encounter variation during their acquisition process, that is to say, even monolinguals and bilinguals
are frequently exposed to dialect due to the intermediate speech repertoire in their community and, hence, societal bidialectism. More importantly, so-called monolingual children may be exposed to asymmetrical language patterns in the home domain from birth onwards. This can be seen from the mother’s and grandparents’ responses regarding their own language choice when speaking to the (grand) child in the home domains. Table 2 illustrates this (cf. Cornips & Hulk 2006):

Table 2. Language choice patterns of participants in Heerlen (Limburg) (n = 30)

<table>
<thead>
<tr>
<th># Child (n = 30)</th>
<th>n = 2</th>
<th>n = 6</th>
<th>n = 8</th>
<th>n = 1</th>
<th>n = 4</th>
<th>n = 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child speaks the local dialect</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Mother speaks dialect in home</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Role relations in home domain:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother dialect to child</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Grandparents dialect to child</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>–</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>

Table 2 reveals which language varieties – Dutch or dialect – are spoken to the thirty examined children at home. This table shows that only 9 children out of 30 can be classified as strictly monolingual: these children speak no dialect at all, and nor do their family members; that is to say, the dialect is not a home language. Only 4 out of the 30 children (13%) can be classified as bidialectal: these children speak dialect
and their mothers also address them in dialect. The majority, however, namely 17 out of 30 children (57%) cannot be classified as monolingual or bidialectal because although the mothers and grandparents speak dialect to their (grand)children in the home domain, the child nevertheless responds to them in Dutch. It is clear that for this type of community, any distinction between monolingual and bidialectal children is meaningless. One of the interesting aspects when it comes to an intermediate speech repertoire is the fact that every child has a passive knowledge of dialect due to the fact that s/he is exposed to dialect in his/her family and community.

The children in Weert can also be classified in monolingual, bidialectal and ‘in-between’ (cf. Dirkx 2012). Table 3 reveals whether these ‘in-betweens’ align with monolinguals rather than bidialectals or vice versa:

Table 3. Overview of 33 children’s ages and PPVT scores (monolingual, bidialectal and ‘in-between’)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>%</th>
<th>PPVT Score</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Range</td>
<td>Mean Range</td>
</tr>
<tr>
<td>Bidialectal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>3</td>
<td>21%</td>
<td>90–91</td>
<td>90</td>
</tr>
<tr>
<td>II</td>
<td>7</td>
<td>50%</td>
<td>98–107</td>
<td>103</td>
</tr>
<tr>
<td>III</td>
<td>4</td>
<td>29%</td>
<td>114–140</td>
<td>125</td>
</tr>
<tr>
<td>In-Between/passive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>0</td>
<td>0</td>
<td>90–91</td>
<td>–</td>
</tr>
<tr>
<td>II</td>
<td>1</td>
<td>17%</td>
<td>102</td>
<td>102</td>
</tr>
<tr>
<td>III</td>
<td>5</td>
<td>83%</td>
<td>109–130</td>
<td>120</td>
</tr>
<tr>
<td>Monolingual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>15%</td>
<td>94–95</td>
<td>95</td>
</tr>
<tr>
<td>II</td>
<td>7</td>
<td>54%</td>
<td>98–107</td>
<td>103</td>
</tr>
<tr>
<td>III</td>
<td>4</td>
<td>31%</td>
<td>113–133</td>
<td>114</td>
</tr>
</tbody>
</table>

Table 3 shows that 83 percent (5 out of 6) of the in-between or passive bidialectal children have to be classified in the most advanced group III, and none of them are in the lowest group I. The split between bidialectals and the ‘in-between’ group reveals that the bidialectals do not differ any more from the monolinguals, as was the case in Table 1 whereas the ‘in-between’ outrank the monolingual control group. One possible reason might be as follows. When they are speaking, the ‘in-betweens’ do not have to split their attention between two language varieties – Dutch and dialect – whereas the bidialectal children do have to do so. In addition, the ‘in-betweens’ are exposed to a much more elaborate range of vocabulary than the monolingual children, and this could lead to them outranking the monolinguals on the PPVT test.

All rights reserved
In sum, 49 children from two primary schools in Weert (Limburg) were classified according to their scores on the standardised Peabody Picture Vocabulary Test. This test is considered as a general indicator of a children’s relative proficiency in Dutch. The bidialectal children differ significantly from the bilingual children in that their test results are much better. The results also show that it is not easy to make a clear-cut distinction between monolingual and bidialectal children in Limburg (or even bilingual children) as we can also distinguish another ‘in-between’ category. The children in this category have mothers who address them in dialect but the children only respond in Dutch. These ‘in-betweens’ or passive bidialectals score the highest on the vocabulary test, and even perform better than the monolingual control group.

Having examined vocabulary, let us now consider how these factors affect the acquisition of syntactic variation, namely the Aux+Inf structure in spoken adult and experimental child Dutch.

3. Aux+Inf and individual variability

The Aux+Inf structure is a good example for showing that (i) children encounter more variation in terms of input than would be expected on the basis of codified grammars and (ii) bidialectal children differ significantly from bilingual children (cf. Cornips 2013b).

The lexical instantiation of Aux+Inf structure in codified Dutch is gaan ‘go’, as in Hij gaat verhuizen ‘He will move’ (see (1) in §1.2. However, there is also a non-standard lexical instantiation in spoken Dutch, namely doen ‘do’ (see (4) below). The absence of do+Inf in codified standard Dutch (as in standard German) is according to Auer (2004:74) “due to a conscious process of purification of the language in which it was purged of its supposedly illogical or superfluous aspects (…)”. After all, the use of do+Inf in declarative sentences can be found throughout the Dutch dialects in the Netherlands with the exception of the provinces Friesland, Groningen and Drenthe (see Barbiers et al. 2008:41) and in spoken Dutch too. One of the reasons to keep do+Inf out of codified Dutch is that it is assumed that this expresses the same meaning as its synthetic equivalent; that is to say, this construction expresses present tense and/or denotes an on-going event. From this perspective, doen ‘do’ is a dummy verb. Although the use of do+Inf is considered dialectal and very marginal in standard Dutch, bilingual adolescents in the standard Dutch area Utrecht nonetheless produce do+Inf in spontaneous speech (cf. Cornips 2013b):

(4) a. dan doe je al denken [Metin, D/T]
   then do you already think
   ‘Then you already think’
Moreover, in dialects and Dutch spoken in Limburg, the use of the declarative do+Inf expresses habitual aspect, as illustrated by the English translation of the Heerlen Dutch example in (5):

(5) ik *doe* timmeren en *opbouwen*
    I do hammer\textsubscript{inf} and build-up\textsubscript{inf}  
    'I am a carpenter and a builder'

However, it seems that not all adult speakers in Heerlen produce the do+Inf structure. The Heerlen corpus consists of 33.5 hours of recorded speech of 67 speakers from the south of Limburg. In this sample, only 18 of the 67 speakers use do+Inf, and according to the independent sociolinguistic variable, most of those speakers are classified as having a ‘low level of education’ (cf. Cornips 1998). The 18 speakers in question produce the construction infrequently (33 times in total). Thus, do+Inf seems to be linked to considerable individual, social and regional variation. Moreover, even the speakers who produce do+Inf do not produce it in all possible cases, so it seems that this is an optional phenomenon.

In acquisition literature, it is also often assumed that in this usage, doen ‘do’ is a dummy verb and is only a device in order to acquire verb second in Dutch (cf. Van Kampen 1997: 50; Zuckerman 2001: 117 and references cited in these studies). In order to test this hypothesis, Zuckerman (2001, cf. also the descriptions in Cornips & Hulk (2005) and Cornips (2013b)) designed a sentence completion test involving 34 picture pairs through which root clauses – V2 – straight order condition – and non-root clauses – OV\textsubscript{FIN} condition – were elicited. He tested 24 children in two age groups: 10 children aged 3;0 to 3;11 and 14 children aged 4;8 to 5;0. Although Aux+Inf was not primed, the results of this experiment show that the youngest children nevertheless produce Aux+Inf. The youngest children (aged three) use Aux+Inf significantly more in root than in non-root clauses (t = 2.583, df = 18, p < .05). The older children, aged five, however, hardly use Aux+Inf. Similar to the adults discussed above, the youngest children show intra- and inter-individual variation. Thus 4 out of 10 children do not use Aux+Inf at all in root clauses whereas the remaining 6 children range in tokens between 1 and 10. Consequently, the children do not show any categorical behaviour with respect to the use of do+Inf, e.g. they never use it in all V2-straight order conditions. Sometimes they use it, other times they do not. Thus, the use of Aux+Inf structure is related to the individual child’s strategy and is optional, similar to the adults’ use of Aux+Inf (cf. Cornips 2013b).
What is more, for that study Zuckerman selected children from bidialectal areas, namely the province of Limburg (south of The Netherlands) and from Groningen (north of The Netherlands). The children from Limburg ($n = 19$) use the lexical auxiliary *doen* in all cases when they produce an Aux+Inf structure whereas the children from Groningen ($n = 5$) who are not being exposed to a dialect exhibiting *doen* ‘do’ use the lexical auxiliary *gaan* ‘go’.

In contrast to children in Groningen and Limburg, bilingual children hardly ever use the Aux+Inf structure. Hulk & Cornips (2005) replicated Zuckerman’s experiment as a pilot involving 9 bilingual children – Dutch/Moroccan, Dutch/Ghanaian, and Dutch/Surinamese – aged between 3;0 and 5;2. The results of the experiment with the youngest bilingual children are complicated and hard to interpret since they display huge individual variation (as in the case of Zuckerman’s monolingual study). One child, Nicole (age 3;6), with a French language background uses Aux+Inf very often both in root (38%) and non-root clauses (48%). The remaining young bilingual children (age 3) differ from the monolingual ones in that they hardly use Aux+Inf. Despite the modelling of a non-root clause in the experiment, both the younger and the older bilingual children use the non-target S-Vf-O word order – in which the lexical finite verb is not in predicate final but in the left periphery of the clause instead of the target S-O-Vf (cf. Hulk & Cornips 2005).\(^1\) Moreover, in a similar experimental task, bilingual English/Dutch children also produce the Aux+Inf structure with an ungrammatical AuxVO order, a structure in which the lexical verb has undergone movement and precedes the object (cf. Cornips 2013b). Overall, bilingual children (Moroccan/Dutch, Surinamese/Dutch, Turkish/Dutch, English/Dutch, French/Dutch) differ from children in bidialectal areas in that (a) they use the Aux+Inf in non-root and root clauses or (b) the youngest children do not use it at all or (c) they construe it in a non-target word order.

It is important to emphasize again that the Aux+Inf structure is always considered to be a dummy one in child language and so-called adult non-standard or dialect varieties whereas in adult standard Dutch, it is considered meaningful – and hence not a dummy verb. However, Aux+Inf in adult Heerlen Dutch is not a dummy verb but rather it expresses habitual aspect. Children who are acquiring their language(s) in Heerlen acquire social and local norms as well. During our fieldwork for the Syntactic Atlas of the Dutch Dialects (SAND, cf. Barbiers et al. 2008; Cornips 2006), an adult speaker in Heerlen tells us explicitly that the use of *do*+Inf is “something that is typical of the southern part of Heerlen. A little bit of

---

1. Note that the non-target S-Vf-O word order is not found among bilingual children by Blom & De Korte (2011) in the results of their almost identical experimental task.
influence from Kerkrade. They add that *doen* ‘do’. The Germans do the same in their dialect’.2

Such local beliefs are revealing since they show that linguistic variation, in this case the (non-)occurrence of *do+Inf*, is socially meaningful. Thus, to a certain extent, there is a process of erasure taking place that simplifies the field in acquisition theory resulting in the variation in spoken speech being overlooked (Irvine 2001). This tendency is very entrenched, and inextricably linked with data-collection. Whereas child data and adult non-standard data involve empirical data, standard data – without any mention of variation – are taken from descriptive (and even prescriptive) grammars.

In any case, it would be interesting in further research to address the question of when children will unlearn the present tense reading of the Aux+Inf structure, and in what kind of conditions and contexts this will happen. A second related question here is the issue of when children acquire aspectual and future readings for *go+Inf* and the habitual reading for *do+Inf*. In sum, the above evidence appears to suggest that there is much more variation and optionality in adult spoken speech and children’s experimental data than might be expected on the basis of an imagined homogenous standard language. Most importantly, children in dialectal areas show a very different output when compared to bilingual children regarding the Aux+Inf structure.

In conclusion, children are confronted with a good deal of variability in their input since the use of *do+Inf* by adults is also an individual optional strategy. The monolingual children themselves also use this as an individual strategy and do so optionally, hence they produce variability themselves. Moreover, the monolinguals in dialectal areas differ significantly from bilinguals in their use of *do+Inf*. Only the youngest children use it in root clauses whereas bilingual children hardly use it, or they use it also in embedded clauses or with a non-target word order.

4. **Bidialectal communities: Grammatical gender of the definite determiner in Dutch**

Dutch has a two-way grammatical gender system which distinguishes between common and neuter gender. This distinction is marked on definite determiners among others. Common nouns take the definite determiner *de* ‘the’, as in *de tafel* ‘the table’, whereas neuter nouns are preceded with *het* ‘the’, as in *het huis* ‘the

---

house’. This gender distinction does not show on the plural definite determiner, which is always *de* ‘the’ and neither does it show on the singular indefinite determiner, which is always *een* ‘a’, as illustrated in Table 4:

<table>
<thead>
<tr>
<th>Definite determiner</th>
<th>Singular</th>
<th>Plural</th>
<th>Indefinite</th>
</tr>
</thead>
<tbody>
<tr>
<td>neuter noun</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>boek ‘book’</td>
<td><em>het</em> ‘the’</td>
<td><em>de</em> ‘the’</td>
<td><em>een</em> ‘a’</td>
</tr>
<tr>
<td>common noun</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tafel ‘table’</td>
<td><em>de</em> ‘the’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are some morphological and semantic regularities in the use of the definite determiner, but these are limited and there are many exceptions (cf. Blom et al. 2008 and references cited there).

In terms of the acquisition of grammatical gender of the definite determiner in Dutch, both experimental and spontaneous speech data show that groups of monolingual (L1) children, of children acquiring two languages from birth (2L1) and also bilingual (L2) children tend to overgeneralise the common determiner *de* ‘the’ with neuter nouns. Monolingual children use *de* ‘the’ instead of *het* ‘the’ until at least age six (Blom et al. 2008). 2L1 and L2 children have been found to continue overgeneralization with *de* ‘the’ even beyond this age (Blom, et al. 2008; Cornips et al. 2006, Cornips & Hulk 2006, 2008; Hulk & Cornips 2006; Unsworth 2008, Unsworth et al. 2011). Roodeenburg & Hulk (2008) attribute the ‘late’ monolingual and bilingual acquisition of Dutch (neuter) gender to a number of internal factors. The lack of a gender distinction in plural and indefinite DPs, the lack of morphological cues on the head noun, the status of *het* ‘the’ as a pronoun and nominalizer in impersonal constructions and with predicative superlatives are all properties of the gender system in Dutch which complicate the learner’s process of discovering unambiguous and salient cues (cf. Unsworth et al. 2011). External factors that are important prerequisites for Dutch gender acquisition for both mono- and bilingual children include the sociolinguistic context (cf. Cornips 2008) and, particularly for bilingual children, the quality of the input (Cornips & Hulk 2008; Unsworth 2008) and longer exposure to Dutch (Unsworth 2008; Unsworth et al. 2011).

In sum, all monolingual children are ‘late’ in acquisition of the grammatical gender of the definite determiner in Dutch. Studies on bilingual children showed that they experience an even longer delay, regardless of whether this feature is present in their first language (as it is in Moroccan-Arabic, Berber, French) or not (as, for example, in Turkish, English, Sranan, see Cornips & Hulk 2008 for
an overview; Unsworth et al. 2011a, b). Cornips (2008) suggests that the extensive delay and variable use of the determiners *de ‘the’ and *het ‘the’ with neuter nouns by Moroccan-Dutch children may be due to social identity construction (cf. Cornips & Hulk 2013).

However, the behaviour of bidialectal children is significantly different from that of bilinguals, especially in terms of the acquisition of grammatical gender in Dutch. Cornips and Hulk (2006) carried out a pilot study on the 30 children discussed above (see Table 2 in §2). In this study, the children were divided into two groups. One was a group of bidialectal speakers (*n* = 13) who were exposed to both the dialect of Heerlen (a Limburg dialect) and Dutch. 9 of these 13 children appeared to be passive bidialectals or ‘in-betweens’ but were classified as bidialectal due to the low number of tokens the children produced and the low number of children involved. The second group consisted of the monolingual children (*n* = 17).

The experiment was a completion test, based on Zuckerman (2001) involving 30 picture pairs (see also §3). In the test, the experimenter presented the child first with a picture with one conjunct and then showed them a second picture with the second conjunct. Then the experimenter presented the pictures again via a coordination structure in which the first conjunct was fully produced by the experimenter and the second conjunct was truncated. The children were asked to complete the sentence and to produce an object, hence a DP (see Hulk & Cornips 2006 for more details). This task is very similar to the one described in Section 3.1 except that (i) the children were asked to produce an attributive adjective and (ii) the sentences were all administered in main clause conditions, as illustrated in (6).3

(6)  

*Experimenter:*

a. Experimenter shows picture of a boy with a green flower  
*Deze jongen tekent de groene bloem en dit meisje tekent …* 
this boy draws the green flower and this girl draws

b. Experimenter shows picture of a girl with a yellow boat  
*… de gele boot* 
… the yellow boat

The results reveal that the bidialectals between 2;0 and 5;0 years old show a developmental pattern in which they have already passed the bare noun stage before  

3. It is important to point out that this type of completion test differs from the one discussed in Section 3.1 since the children do not repeat a DP introduced by the experimenter. For this reason, the data from the bilingual children and dialect children and the monolinguals as controls cannot be compared directly since they are the results of two different experiments.
the age of 3;0 whereas the monolinguals only do so between 3;0 and 5;0 years of age. Moreover, the monolingual children between 3;0 and 5;0 years show no progress in the acquisition of the neuter *het*, whereas the bidialectals in the same age group do. There is thus a clear trend showing that bidialectal children are *ahead* of monolingual controls which is very different from what is found regarding bilingual children.

Cornips and Hulk (2008) suggest that the gender system of the dialect grammar must be one of the factors contributing to the progress of bidialectal children. Dialects in Limburg, in contrast to Dutch make (a) a three-way distinction between masculine, feminine and neuter nouns and (b) this distinction is morphologically visible on both the indefinite and definite determiner, as illustrated in Table 5:

Table 5. The morphology of determiner in the dialect of Heerlen (Limburg) (cf. Cornips 1998)

<table>
<thead>
<tr>
<th>Determiner</th>
<th>Definite</th>
<th>Indefinite</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEUTER</td>
<td>'t</td>
<td>e</td>
<td></td>
</tr>
<tr>
<td>MASCULINE</td>
<td>d'rne</td>
<td>inne</td>
<td>de</td>
</tr>
<tr>
<td>FEMININE</td>
<td>de</td>
<td>ining</td>
<td></td>
</tr>
</tbody>
</table>

The three-way distinction in the dialects, and the morphological distinction on the indefinite determiner mean that a child who speaks dialect is able to acquire more salient cues for a gender feature. What is more, the (structural) similarity of the gender systems in the two varieties may reinforce the bidialectal child’s awareness of grammatical gender in Dutch and consequently increase their rate of acquisition. Having examined the aspect of gender it is now interesting to consider how this relates to a new dataset from Limburg.

4.2 New study: Monolingual, bidialectal and bilingual children

In order to examine whether we could reduplicate our results from the 2006 Cornips and Hulk study, the 49 children in Weert (see Table 1&3 in §2) were tested not only on vocabulary but also participated in a picture description task and a Story task about grammatical gender of the definite determiner in Dutch (see Unsworth et al. (2011a, b). Here, the bidialectal children (n = 20) are compared to the bilingual children (n = 16) in order to find more evidence to support the claim that bidialectal acquisition is of a different type to bilingual acquisition.

The children were classified as bidialectal (n = 20) if they had been exposed to dialect and Dutch from birth onwards; in addition, 16 children were classified
as bilingual children. They have the following languages: Moroccan Arabic/Berber (n = 11), Turkish (n = 2), Polish (n = 1), Dari (n = 1) and Rumanian (n = 1).

The children’s knowledge of grammatical gender marking on definite determiners was tested using two elicited production tasks, following Blom, Polišenskà & Weerman (2008), and also, especially following Unsworth & Hulk (2010) and Unsworth et al. (2011a, b). The general set-up of the picture description task is that the children are presented with pictures of the nouns in question on a computer screen and first asked to name them, thereby eliciting an indefinite noun. Subsequently, they were asked a question about the same object (e.g. “Which object is brown?”) or prompted to describe the position of another object relative to the object of interest (e.g. “The ball is in front of … (child: … the (yellow) robot”), thereby eliciting a definite determiner in either a simple (Det-N) or complex (Det-Adj-N) DP. This procedure and tests are an exact repetition of the ones used by Unsworth and Hulk (2010). Each noun was elicited once in a simple DP (det-N) and twice in a complex DP (det-Adj-N). In both tests, the participants were given 12 items, including 6 neuter and 6 common ones.

In the Story task, the experimenter explains that he or she will tell the child three short stories, and that the child should help with this task. The child is expected to complement a sentence presented by the interviewer – an asterisk in (7) below – while a PowerPoint shows the relevant pictures (see Unsworth 2008; Unsworth & Hulk 2010).


“John and Mary are brother and sister. One day, they are going out to have some fun. They visit an animal farm for children. There they see a [*] [sheep], a [*] deer and a [*] rabbit. The animals are hungry. Jan and Mary give them something to eat [*]. Which of these animals get a sandwich? [the deer] [*]…”

Let us now consider the results for the simple DP’s from the bidialectals and bilinguals to see whether bidialectals differ from bilinguals as we might predict from the vocabulary scores.

Table 6 reveals that all bidialectal groups show a target response for the common definite determiner de ‘the’, whereas this is only true for the bilinguals’ in Groups II and III. The bidialectals differ significantly from the bilinguals in that there is a significant development (Fisher Exact p = 0,03) between Groups II and III in the production of the target neuter determiner. The bilinguals do not show any development with respect to the target production of neuter het ‘the’. The
The four monolingual children in Group III scored 25% lower than the bidialectal children in their target use of neuter definite determiner *het* ‘the’.

Table 7 contains the results for the bidialectal and bilingual children’s target production of the grammatical gender of the definite determiner in complex DP’s containing an adjective:

The results in Table 7 are similar to those in Table 6. All bidialectal groups show a target response for the common definite determiner *de* ‘the’, whereas this is only true for the bilinguals in Group III. One crucial difference between the bidialectals and the bilinguals is that the former show a significant development in the production of the target neuter determiner between Groups I and II. In contrast, the bilinguals do not show any development with respect to the target production of neuter *het* ‘the’. The four monolingual children in Group III scored significantly lower with a difference of 39% than the bidialectal children on the target use of the neuter definite determiner *het* ‘the’. Furthermore, the monolinguals in Group III

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>PPVT</th>
<th>Simple DP’s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>de</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>neuter</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18/22 82%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>19/19 100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25/47 53%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>47/48 98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>14/43 33%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50/50 100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>32/48 67%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>39/56 70%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>14/22 64%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>22/22 100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4/4 100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4/4 100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11/19 58%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>22/23 96%</td>
</tr>
</tbody>
</table>
score higher in simple DP’s than complex DP’s, and the same holds for the bidialectals in Group I. The bilinguals show more mixed results.

In sum, it is clear that bidialectal acquisition is significantly different from bilingual acquisition with respect to the grammatical gender of the neuter definite determiner in both simple and complex DP’s. Bidialectal children show significant developments whereas there is no such development for bilingual children. Moreover, bidialectal children in the most advanced group are far ahead of monolingual children: 67 percent of their responses are target results whereas the bilinguals score only 10 percent of target results.

5. Conclusion

It is extremely important to take into consideration the socio-syntactical perspective in acquisition research and, moreover, to investigate whether children’s use of so-called non-target or ungrammatical forms in the standard language may
be a reflection of the fact that these forms are socially meaningful for the child and her community and not only a reflection of the child’s lack of grammatical competence or incomplete/unsuccesful acquisition. Indexical sensitivities may account for how and why children might not produce standard grammatical forms in situated contexts although their teachers and/or parents want them to speak the standard language. A variable child language output is thus not the outcome of developmental stages per se since the outcome may reflect and construct social identities in the local communities.

The aim of this paper was two-fold. First, it is argued that the input to which children are exposed is inherently variable due to both language internal and also language external factors as well. While syntax is often viewed within sociolinguistics as a marker of cohesion in large geographical areas, syntactic variants may also act as marker of local identity. The link between social and linguistic elements has to be understood as very complex and multi-dimensional since language choice and linguistic forms can index membership to various social and local groups. This socio-syntax perspective therefore by definition suggests that the configuration of the input is very complex. The Aux+Inf construction in spoken Dutch was taken as an example to show that speakers reveal much more variation than one would assume on the basis of the standard grammar. This type of construction may express a habitual reading which goes unnoticed in standard codified grammars. Adults use the Aux+Inf construction as an individual strategy and the phenomenon is optional. This individual strategy and optionality is also seen to be present when we examine the experimental data of children acquiring Dutch.

The second aim of the paper was to address the question of whether bidialectal acquisition is the same type as bilingual acquisition. The answer here is negative. Bidialectal children score significantly higher on vocabulary than bilingual speakers, as well as using the Aux+Inf structure differently and also acquiring neuter gender of the definite determiner significantly faster and with more ease than bilingual children.

Most importantly, in this paper the distinction between monolingual and bidialectal children in so-called bidialectal areas is problematized. Many children grow up in households in which the principal caregivers address them in dialect but the child uses Dutch only. These children with passive knowledge of dialect can be termed ‘in-between’, and they achieve vocabulary scores which are better than those of monolingual children. Bidialectal children also outrank monolingual children in the acquisition of the neuter definite determiner in Dutch. The question now arises of whether in the context of bidialectal communities we can maintain that a standard language and dialect are discrete varieties, and two independent targets in the acquisition process. The answer to this
question is no if the speech repertoire in these areas is an intermediate one in which there exists a very close structural relationship between the standard variety and dialects.

References


Blommaert, Jan. 2010. The sociolinguistics of globalization. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511845307


All rights reserved


**Author’s address**

Leonie Cornips  
Meertens Instituut (KNAW)  
& Maastricht University  
Joan Muyskenweg 25  
P.O. Box 94264  
1090 GG Amsterdam  
The Netherlands  
leonie.cornips@meertens.knaw.nl