Late language acquisition and identity construction

Variation in use of the Dutch definite determiners de and het*

Leonie Cornips & Aafke Hulk
Meertens Institute & Maastricht University / University of Amsterdam

Acquisition of the grammatical gender of the Dutch definite common determiner de and neuter het is a long-lasting process since monolingual children do not acquire a target grammar with respect to the use of het until the age of six. Before that age, they overuse de. Bilingual child acquirers from ethnic minority communities show an overuse of de to a much higher extent than their monolingual controls. We will explore the relation between language acquisition and language variation. It is argued that when the acquisition process of a grammatical phenomenon takes too long, it will become vulnerable in the sense that language external and internal factors start to interfere with this process and the emerging variation will remain for some acquirers. The hypothesis is put forward that in that case a grammatical phenomenon is very eligible to be used in identity construction.

1. Introduction

In acquisition studies, it is assumed that the child is cognitively equipped such that she explores the linguistic possibilities within a specific language and stabilizes on a language that is target-like, i.e. equivalent to that of the adults in her linguistic community. In this paper, it is argued that when this exploration takes too long for monolingual children, as is the case with grammatical gender of the definite determiner in Dutch, language external and language internal factors start to interfere with this process and one may expect to find variation to arise in the language

* The acquisitional part of this research has been conducted in collaboration with Claasje Reijers and Paz González and has been presented in Utrecht and Lissabon (cf. Cornips et al. 2012; Reijers et al. 2009).
acquired by children (Cornips & Hulk 2008). Regarding external factors, this paper puts forward the hypothesis that grammatical categories that are acquired late are the ones that are most eligible in identity construction in urban youthful speech practices by speakers having two languages at their disposal. It is extremely important to investigate this hypothesis since the child’s use of so-called ungrammatical forms such as the Dutch definite determiner *de* instead of *het* may be a reflection of its social meaningfulness and not a reflection of the child’s lack of grammatical competence or unsuccessful bilingual acquisition of standard Dutch (Cornips 2008).

Dutch makes a difference between neuter and common only for singular, definite nouns, as illustrated in Figure 1.

Figure 1. The singular and plural definite determiners in Dutch

<table>
<thead>
<tr>
<th>Gender</th>
<th>Singular definite determiner</th>
<th>Plural definite determiner</th>
<th>Singular indefinite determiner</th>
<th>Plural indefinite determiner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuter noun</td>
<td><em>boek</em> “book”</td>
<td>determiner <em>het</em> “the”</td>
<td>determiner <em>de</em> “the”</td>
<td>determiner <em>een</em> “a”</td>
</tr>
<tr>
<td>Common noun</td>
<td><em>tafel</em> “table”</td>
<td>determiner <em>de</em> “the”</td>
<td>determiner <em>een</em> “a”</td>
<td></td>
</tr>
</tbody>
</table>

Regarding the actuation problem, there is not much evidence for grammatical gender in the input in Dutch, in contrast to Italian, French and/or German (cf. Pérez-Pereira 1991; López-Ornat 2003). Unlike these languages, Dutch has hardly any phonological cues on the noun for common or neuter. In addition, the distinction between neuter and common is almost arbitrary. Moreover, grammatical gender is visible only on single definite determiners, i.e. common *de* and neuter *het* and demonstratives, i.e. common *die/deze* “this” and neuter *dat/dit*. So, indefinite and plural determiners show no morphological marking regarding the gender of the noun they combine with (cf. Figure 1). Finally, common nouns which take *de* outnumber neuter nouns taking *het* by a ratio of about 2:1 regarding type frequencies and by a ratio 3:1 when based on token frequencies (Blom et al. 2008: 302). It is probably due to this scarce evidence of gender that monolingual acquisition of the Dutch definite determiners is a long-lasting process, i.e. children do not acquire a target grammar until the age of six with respect to the target-like use of *het*. Until that age, they overgeneralize the definite determiner *de* where *het* is required (Blom et al. 2008). For bilingual children, this process takes even more time (Cornips & Hulk 2008; Unsworth et al. 2011).

Roodenburg and Hulk (2008) and Cornips and Hulk (2008) suggest that when children in Dutch start using definite determiners it is not a gender marker (yet) but
the semantic property \([\pm \text{count}]\) plays a role in the choice of the definite determiner. More specifically, they examine the hypothesis put forward by Picallo (2008) that \textit{neuter} pronouns in languages like Spanish do not represent a third gender class – next to feminine and masculine pronouns – but instead are elements that lack grammatical gender and number. In this paper, we take this assumption further. When children start using \textit{het}, they consider this determiner to be an element that agrees with the \([-\text{count}]\) characteristics of a noun. The first aim of this paper is to test this assumption in young monolingual Dutch and Spanish-Dutch bilingual children (Reijers et al. 2009; Cornips et al. 2012). The second aim is to discuss the social meaning of the overgeneralisation of \textit{de} by bilingual adolescents.

Since both monolingual and, in particular, bilingual children from ethnic minority communities show problems acquiring the definite determiner \textit{het}, it cannot just be external factors, such as the quantity and the quality of the input or identity construction, that explain difficulties in the late acquisition of grammatical gender in Dutch. Internal, linguistic factors must play a role, too. According to Sorace (2005) external interfaces between syntax and other cognitive systems and internal interfaces among connections between different components of the grammar are predominant loci of instability. The acquisition of the Dutch determiners and their gender involves interface domains, e.g. connections between several components of grammar, namely lexicon, morphology and syntax. What is needed for their acquisition is (i) the presence of the syntactic position D, (ii) the lexical and grammatical features of the noun, (iii) agreement between D and N, and (iv) the features and morphology of the determiner. Thus, the acquisition of grammatical gender involves more than only lexical learning.

This paper is organized as follows. In Section 2 we introduce the hypothesis, the subjects and methodology of the experiment. In Section 3, we discuss the construction of a ‘street-wise’ identity through \textit{de} by adolescents of ethnic minority communities. The last section will be devoted to a conclusion.

### 2. Linguistic factor: Count/mass distinctions in nouns and determiners

Dutch allows singular nouns to be bare if they are \([-\text{count}]\):

\[
\begin{align*}
(1) & \quad \emptyset \text{ melk} \quad \text{\textit{een}} \text{ melk} \quad \text{\textit{melk-}\textsubscript{PL}} \quad \text{“milk”} \quad \text{[–count]} \\
& \quad \emptyset \text{ water} \quad \text{\textit{een}} \text{ water} \quad \text{\textit{water-}\textsubscript{PL}} \quad \text{“water”} \quad \text{[–count]}
\end{align*}
\]

Moreover, only singular \([+\text{count}]\) nouns take the indefinite article \textit{een} “a”, and only count nouns can be pluralized in Dutch:

\[
\begin{align*}
(2) & \quad \text{\textit{\textbullet \ tafel}} \quad \textit{een} \text{ tafel} \quad \text{\textit{tafel-}\textsubscript{PL}} \quad \text{“table”} \quad \text{[count]} \\
& \quad \text{\textit{\textbullet \ boek}} \quad \textit{een} \text{ boek} \quad \text{\textit{boek-}\textsubscript{PL}} \quad \text{“book”} \quad \text{[count]}
\end{align*}
\]
This implies that there is clear evidence in the input to children for a classification of Dutch nouns in [count] versus [–count], more so than for [neuter] versus [common]. If Dutch children initially use this classification to select *de* and *het*, they may assume that:

a. *de*-nouns that can be pluralized and combined with *een* in the singular, select *de*, i.e. nouns with the semantic feature [+count];

b. *het*-nouns that cannot be pluralized and do not take *een* in the singular, select *het*, i.e. nouns with the semantic feature [–count] take *het*.

These hypotheses predict that children will initially use *het* only with [–count] nouns such as *water* (following Picallo 2008), and not with [+count] nouns such as *boek* “book”.

2.1 Methodology

In order to examine whether monolingual and bilingual i.e. Spanish-Dutch children combine the determiner *het* initially most often with [–count] nouns and only later with [+count] and/or [+neuter] nouns, we selected 21 monolingual and 17 bilingual children (total of 38) in two monolingual age groups between 4;2-4;11 and 5;1-5;11 and two bilingual groups between 3;6-4;8 and 5;1-6;7 years. All bilingual children have Dutch as one of their home languages.

A sentence completion test, describing 34 picture-pairs (cf. Cornips & Hulk 2008) enables us to investigate the use of definite determiners with respect to [common], [±count] and [neuter], [±count] test items. For each phenomenon, the subjects were asked to use 22 nouns, divided into 12 [+count, singular] nouns (6 common and 6 neuter nouns) and 10 [–count] nouns (6 neuter and 4 common nouns). We also investigated 12 [+count, plural] nouns in the condition [neuter] (n = 6) and [common] (n = 6).

The investigator introduces two pictures consecutively and asks the child to complete the sentence relating to the object shown in the second picture. The test format requires the child to complete the sentence, as illustrated in (3) and (4):

(3) Investigator: *Dit is een konijn en dit is een schaap. Dus dit is het meisje met het konijn en dit is het meisje met…?*

“This is a rabbit and this is a sheep. So, this is the girl with the rabbit and this is the girl with …?”

Child: *de schaap* (schaap “sheep” is [+count]; expected answer according to a. above is *de*)

“the sheep”
(4) Investigator: *Dit is een bed en dit is zand. Dus dit is het meisje in het bed en dit is het meisje in …?* “This is a bed and this is sand. So, this is the girl in the bed and this is the girl in …?”

Child: *het zand* (*zand* “sand” is [-count]; expected answer according to b. above is *het*) “the sand”

2.2 Results of the monolingual children and discussion

The hypotheses we adopted above predict that when Dutch children use *het* they use it (most) with [-count] nouns:

(5) a. *het water*$_{N/-CNT}$ “water” but also
b. *het sneeuw*$_{C/-CNT}$ “snow”, but not (yet)
c. *het paard*$_{N/+CNT}$ “horse”

In contrast, overgeneralization with *de* will be then restricted to [+count] nouns:

(6) a. *de boek*$_{N/+CNT}$ “book” and
b. *de paard*$_{N/+CNT}$ “horse” but not:
c. *de water*$_{N/-CNT}$ “water”

The results of the monolingual children regarding the use of the definite determiners *de* and *het* for the [+count] and [-count] nouns are presented in Table 1 and Table 2, respectively.¹

Table 1. Monolingual results for the (in)definite determiner with [+count] nouns with respect to the conditions ‘neuter’ and ‘common’

<table>
<thead>
<tr>
<th>Age</th>
<th>[+count, neuter, +singular]</th>
<th>[+count, common, +singular]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>de</em></td>
<td><em>het</em></td>
</tr>
<tr>
<td>4;2-4;11</td>
<td>33.3%</td>
<td>14.8%</td>
</tr>
<tr>
<td>n = 9</td>
<td>18/54</td>
<td>8/54</td>
</tr>
<tr>
<td>5;1-5;11</td>
<td>27.8%</td>
<td>34.7%</td>
</tr>
<tr>
<td>n = 12</td>
<td>20/72</td>
<td>25/72</td>
</tr>
</tbody>
</table>

*het boek* “the book”

*de bal* “the ball”

¹ Although a category bare noun and ‘other’, that is to say, answers other than *de*, *het*, and *een* have been calculated, it is not presented in the tables.
Table 2. Monolingual results for the (in)definite determiner with \([-\text{count}]\) nouns with respect to the conditions ‘neuter’ and ‘common’

<table>
<thead>
<tr>
<th>Age</th>
<th>([-\text{count}, \text{neuter}])</th>
<th>([-\text{count}, \text{common}])</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(de)</td>
<td>(het)</td>
</tr>
<tr>
<td>4;2-4;11</td>
<td>25.9%</td>
<td>24.1%</td>
</tr>
<tr>
<td>(n = 9)</td>
<td>14/54</td>
<td>13/54</td>
</tr>
<tr>
<td>5;1-5;11</td>
<td>8.3%</td>
<td>55.6%</td>
</tr>
<tr>
<td>(n = 12)</td>
<td>6/72</td>
<td>40/72</td>
</tr>
</tbody>
</table>

First, in general, the children produce much more correct \(de\) (63% for all common nouns between 4;2-4;11, 76% between 5;1-5;11) than \(het\) (19% for all neuter nouns between 4;2-4;11, 45% between 5;1-5;11). These percentages of target-like \(het\) are lower compared to the monolingual experimental results of Blom et al. (2008) in which children between 4;0-4;11 \((n = 17)\) and between 5;0-5;11 \((n = 15)\) produce target-like \(het\) in 44% and 69% of the cases, respectively. Second, the children overuse \(de\) considerably with neuter but hardly overuse \(het\) with common nouns.

Most importantly, the children use \(de\) and \(het\) both with [+count] and [–count] nouns. We did not expect them to use \(de\) with a [–count] noun such as \(sneeuv\) “snow” but they did. We also did not expect them to use \(het\) with a [+count] noun such as \(boek\) “book” but they did. Therefore, these results do not support the hypotheses proposed above, i.e. that children would (initially) use \(het\) with [–count] nouns and \(de\) with [+count] nouns. Nevertheless, within the group of neuter nouns, both the 4 and 5 year olds use \(het\) significantly more with [–count] nouns such as \(gras\) “grass” than with [+count] nouns, such as \(boek\) “book”. Fisher’s Exact Test reveals that the distribution for \(het\) with [–count] nouns significantly differs from that of \(het\) with [+count] and from \(de\) with [+count] nouns. These interactions are significant for both the four and five year old children (Fisher’s Exact Test \(p < .0001\)). Thus, the cue [–count] does play a significant role in the selection of \(het\), but it operates in interaction with the grammatical gender feature of the noun.

What is more, although we aimed to elicit definite determiners, some children produced indefinites, using \(een\), but they use it only with [+count] nouns, not with [–count] nouns (only 1 out of 210 tokens) and never with plurals. This suggests that both the four and the five year olds know the difference between [+count] and [–count] nouns. This knowledge is a prerequisite for the possible use of the cue [–count] for other purposes, such as the selection of \(het\) as definite determiner. However, we have to keep in mind that this is only a study with low numbers of children, so we have to be cautious in generalizing the results.
2.3 Results of the bilingual children and discussion

Tables 3 and 4 present the results of the bilingual Spanish-Dutch children for the [+count] and [–count] nouns, respectively.

Table 3. Bilingual results for the (in)definite determiner with [+count] nouns with respect to the conditions ‘neuter’ and ‘common’

<table>
<thead>
<tr>
<th>Age</th>
<th>[+count, neuter, +singular]</th>
<th>[+count, common, +singular]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>de</td>
<td>het</td>
</tr>
<tr>
<td>3;6-4;8</td>
<td>23.8%</td>
<td>7.1%</td>
</tr>
<tr>
<td>n = 7</td>
<td>10/42</td>
<td>3/42</td>
</tr>
<tr>
<td>5;1-6;7</td>
<td>28.3%</td>
<td>23.3%</td>
</tr>
<tr>
<td>n = 10</td>
<td>17/60</td>
<td>14/60</td>
</tr>
</tbody>
</table>

**Table 4.** Bilingual results for the (in)definite determiner with [–count] nouns with respect to the conditions ‘neuter’ and ‘common’

<table>
<thead>
<tr>
<th>Age</th>
<th>[–count, neuter]</th>
<th>[–count, common]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>de</td>
<td>het</td>
</tr>
<tr>
<td>3;6-4;8</td>
<td>14.3%</td>
<td>16.7%</td>
</tr>
<tr>
<td>n = 7</td>
<td>6/42</td>
<td>7/42</td>
</tr>
<tr>
<td>5;1-6;7</td>
<td>18.3%</td>
<td>50.0%</td>
</tr>
<tr>
<td>n = 10</td>
<td>11/60</td>
<td>30/60</td>
</tr>
</tbody>
</table>

Again, in general, the bilingual children produce much more correct *de* (34% for all common nouns between 4;2-4;11, 72% between 5;1-5;11) than *het* (24% for all neuter nouns between 4;2-4;11, 37% between 5;1-5;11) and overuse *de* with neuter but not *het* with common nouns. The bilingual children are also similar to the monolinguals in that they use *de* and *het* both with [+count] and [–count] nouns. This similarly contradicts the hypotheses proposed above. However, just like their monolingual peers, all bilingual children use *het* more often with [–count, +neuter] nouns such as *gras* “grass” than with [+count, +neuter] nouns such as *boek* “book” (Fisher Exact Test, age 3;6-4;8 p < .0001, age 5;1-6;7 p < .0014). This is clear evidence that at these ages the bilingual children also use the feature [–count] in their selection of *het*. 
Just like the monolinguals, the bilinguals sometimes produce the indefinite determiner *een* instead of the definite one but they do so only with [+count] nouns and never with [–count] showing that they indeed know the difference between [+count] and [–count] nouns.

In brief, we have seen that the bilingual Spanish-Dutch children show a similar acquisition pattern as their monolingual peers, and use both the gender and the [+count] features of the noun in their use of definite determiners. This linguistic factor may explain the late target-like use of *het* by both monolingual and bilingual children. These children are rather young. In the next section, we will discuss older bilingual speakers of minority ethnic minority communities in order to investigate the social meaning of overuse of *de* instead of *het*.

3. **Spontaneous speech data and identity work**

Previous experimental research shows that bilingual children speaking Berber-Dutch and Turkish-Dutch reveal problems in using target-like the definite determiner *het* (cf. Cornips & Hulk 2008). In one specific study, 24 Turkish and Berber-Dutch speaking children, aged between 10;5 through 12;11 overuse *de* at that age still in 49% of the cases whereas they use *het* target-like in 42% of the cases (Cornips et al. 2006). Apparently, the children know that Dutch nouns fall into two classes. There is no experimental data available of adolescent speakers but at the end of the 1990s, a considerable amount of spontaneous speech of youngsters having two languages at their disposal, i.e. Dutch/Moroccan, /Turkish and /Hindi, was collected in the Utrecht neighbourhood Lombok/Transvaal (approximately 8 hours). All speakers were in their early twenties. All youngsters, regardless of their other languages, reveal a tendency to use the definite determiner *de* in cases where the definite determiner *het* is required. In all these cases, the noun has the property [+count], as illustrated in (7) through (9). It is important to point out that it is only with respect to grammatical gender that these speakers differ from monolingual Dutch speakers, that is to say, they are standardlike regarding other phenomena in Dutch such as negation, expletive constructions and word order (Cornips 2008).

(7)  

<table>
<thead>
<tr>
<th>a.</th>
<th><em>de</em> geloofN+CNT</th>
<th><em>Hassan, Turkish/Dutch</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>the faith</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td><em>de</em> hele gebouwN+CNT</td>
<td><em>Guray, Turkish/Dutch</em></td>
</tr>
<tr>
<td></td>
<td>the entire building</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td><em>de</em> laatste jaarN+CNT</td>
<td><em>Hassan, Turkish/Dutch</em></td>
</tr>
<tr>
<td></td>
<td>the last year</td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td><em>de</em> boekN+CNT</td>
<td><em>Badir, Moroccan/Dutch</em></td>
</tr>
<tr>
<td></td>
<td>the book</td>
<td></td>
</tr>
</tbody>
</table>
e. *de kanaal* \(_{N/+CNT}^{N/+CNT}\) the channel \[\text{[Naraen, Hindi/Dutch]}\]

Also singular diminutives – although they are always neuter but [+count] in standard Dutch – may be combined with *de* instead of required *het*:

(8) a. *de filmpje* \(_{N/+CNT}^{N/+CNT}\) the movie \[\text{[Badir, Moroccan/Dutch]}\]

b. *de vaste groepje* \(_{N/+CNT}^{N/+CNT}\) the steady group \[\text{[Youssef, Turkish/Dutch]}\]

In accordance with the experimental results described above, the youngsters use *het* in two cases where *de* should be used. In both cases the noun has the property [–count]:

(9) a. *het nadruk* \(_{C/-CNT}^{C/-CNT}\) the emphasis \[\text{[Guray, Turkish/Dutch]}\]

b. *het muziek* \(_{C/-CNT}^{C/-CNT}\) the music \[\text{[Badir, Moroccan/Dutch]}\]

Young native-born descendents of immigrants meet each other in very diverse urban contexts – as do other newcomers – creating a linguistic diversity or using linguistic variants which are not necessarily grammatical in the dominant variety but are there for purposes of creating alternative solidarities (Rampton 1995). We argue that the variable use of *de* and *het* plays a role in identity construction as well (Cornips 2008). Let us consider the following interview with the adolescent Moroccan speaker S. in Rotterdam (Nortier & Dorleijn 2008: 132):

S: That is the bad kind of Dutch.
J: Does it have a name?
S: No, not really, but in principle you uhmm.. just use the articles deliberately in the wrong way.
M: Right! So you use them in the wrong way deliberately? Just like-
S: – Yes, like
J: *Die meisje* (“that girl” – *dat meisje* in standard Dutch)
S: I would say: *Die huis* (“that house” – *dat huis* in standard Dutch). At the same time I know, I mean, I am very well aware of the fact that it should actually be *het huis*, but it would make a dumb impression if I would say....
M: Yes
S: If I would say *dat huis* in the street
M: Yes, yes.
S: It is just *die huis*. But when I speak with you two (the authors both Dutch and middle-aged JN/MD) it is just *dat huis*.
Nortier and Dorleijn (2008: 132) claim: “The speaker in the quotation explicitly says that he has to make errors, deviations from the standard norm, in order to be recognized as someone who is hanging out with friends”. Importantly, they point out that youngsters of Moroccan descent are developing both linguistic norms and norms of stylistic appropriateness. We like to argue that in doing so the overuse of *de* is crucially involved. Identity construction i.e. “being in the streets with friends” takes place through the use of the definite determiner *de* and demonstrative determiner *die* instead of target-like *het* and *dat*. Consequently, the overgeneralization of *de* in this group of children cannot be ascribed to bilingual acquisition effects alone, but is also governed by linguistic norms of the members of a group (Le Page & Tabouret-Keller 1985). Their overuse is probably a case of almost reaching a proficiency in gender marking as monolinguals. And on hitting adolescence they chose to imitate what was initially an L2 feature which itself has been reallocated to an identity function (cf. Cornips 2008).2

4. Conclusion

In this paper we have explored the relation between language acquisition and language variation. It is shown that target-like acquisition of grammatical gender in Dutch, i.e. the selection of the definite determiner *het* is late due to a linguistic factor, namely the interplay of the semantic feature [±count] with the syntactic gender feature. This interplay results in a situation of long-lasting variation in use of both the definite determiners *de* and *het* with the same noun. It is proposed that also language external factors start to play a substantial role resulting in the ‘maintenance’ of this type of language variation since the acquisition of grammatical gender in Dutch takes too long. Children having two languages at their disposal who are growing up in ethnic minority communities, in particular, go through a long-lasting acquisition process of grammatical gender. It is discussed that the overuse of *de* is put forward by speakers growing up in ethnic minority communities in the construction of a street-wise identity “hanging out with friends”. These findings are crucial since the use of so-called ungrammatical forms may be a reflection of their social meaningfulness and not a reflection of an unambiguous unsuccessful bilingual acquisition of standard Dutch only.

2. Thanks to the anonymous reviewer who has pointed this out.

© 2013, John Benjamins Publishing Company
All rights reserved
References


