THE ACQUISITION OF GRAMMATICAL GENDER IN WELSH

Virginia C. Mueller Gathercole
Enlli Môn Thomas
University of Wales Bangor
University of Wales Bangor

Nadine Laporte
University of Wales Bangor

Two experiments examine the development of grammatical gender in Welsh-speaking children. Experiment 1 explores children's productive abilities with local cues to gender, Experiment 2 children's interpretative abilities with long-distance agreement for gender. Participants in Experiment 1 were children between 4;0 and 9;7 and in Experiment 2 between 4;8 and 11;1. All were from North Wales. In Experiment 1, subjects came from Welsh-only homes; in Experiment 2, they came from Welsh-only, English-only, or Welsh-English homes. Results suggest that when a language has a complex gender system that is marked by opaque morphophonological processes, the course of development is protracted and variable. Children acquire Welsh gender in relation to noun marking, adjective marking, and long-distance agreement in a piecemeal fashion. Use for human referents appears most productive, but extension of the system beyond use for human referents seems to occur to some degree.

Introduction

One of the characteristic features of Celtic languages is the pervasive presence of mutations, morphophonological changes that affect the initial consonants of words, but which are triggered primarily by morphosyntactic conditioning. In Welsh, there are three major mutation types, Soft Mutation, Nasal Mutation, and Aspirate Mutation, each of which occurs in a variety of lexical and syntactic contexts. The primary goal of this study is:

1. This research was supported in part by ESRC grant R000 23 7882 to the first author and by a Bookland and Development Trust studentship to the second author. We wish to gratefully acknowledge the untiring and very generous help of Dafydd Idriswyn Roberts, Carys Jones Gwyn Howelis, Rhian Williams, Delyn Waters, Siwan Jones, Non Hughes, Olwen Green Janet Jones, and the teachers, parents, and children at Ysgol Gynradd Llanfairpwllgwyngyll at Ysgol Y Gaernedd, at Ysgol Y Gelli, at Meithrin, at Ysgol Gynradd Capel Coch, at Ysgol Parc y Bont, at Ysgol Gynradd Y Felinheli, at Ysgol Gynradd Tregarth, and at Ysgol Gynradd Abercynon. We would also like to thank Colin Baker, Margaret Deuchar, and especially Norma Roberts, for their assistance in the preparation and administration of these studies.

Journal of Celtic Language Learning 6 (2001), pp. 53-87
to examine children's productive and receptive command of soft Mutation in relation to one of these contexts, grammatical gender. The study will also, secondarily, provide information on the acquisition of Aspirate Mutation in relation to gender.

Research on the acquisition of grammatical gender across languages has shown that for many languages, children gain an early command of gender. Children have been shown in a number of languages to construct or learn the systems largely on the basis of distributional privileges (Popova 1973; Karmiloff-Smith 1978, 1979; Levy 1983a, 1983b, 1996; Berman 1985; Smoczyńska 1985; Mills 1986; Cain, Weber-Olsen, and Smith 1987) and to do so swiftly and effortlessly (e.g. Smoczyńska 1985). Often in the languages studied, however, gender marking is quite overt and provides a clear one-to-one correspondence between a marker and the gender encoded. When the system is somewhat more opaque, the acquisition of gender appears more difficult and more protracted (e.g. Maratos and Chalkley 1980; Mulford 1985; Smoczyńska 1985), as is often the case with opaque systems (Lieven 1997). In fact, in such cases, children may make use of semantic information to help them break into the system (Mulford 1985).

Gender marking in Welsh is primarily encoded as part of a complex system involving Soft Mutation (henceforth ‘SM’), a process of lenition applying to stops, /m/, and liquids. There are multiple form-function mappings between SM and what it encodes, which makes the grammatical gender system quite opaque. Because the system is complex, it allows us to address a number of theoretical questions regarding the acquisition of gender and regarding the acquisition of language in general:

1. How do children go about acquiring a complex/opaque grammatical gender system in their language?

2. To what extent do children rely on a variety of clues to categorization in acquiring gender constructs?

3. To what extent do children construct abstract representations of the grammatical system they are learning, or, conversely, to what extent does their knowledge remain more piecemeal and closer to the surface?

4. How does children’s performance on local marking of gender compare with their distant marking of gender?

5. To what extent does the amount of exposure to the Welsh language affect the answers to these questions?

The two experiments reported here attempt to provide some answers to these questions.

**Mutation and the grammatical gender system**

Under SM, word-initial voiceless stops and liquids (i.e. /p, t, k, ñ, ñ/) become voiced (/b, d, g, l, r/), /m/ and the voiced stops /b, d/ become fricatives (/v, ɣ, θ/), and /g/ gets deleted. The conditioning environments for the occurrence of SM are numerous. Table 1 below presents some samples of environments that trigger SM; these include occurrence immediately after certain prepositions, after certain possessive forms, after certain conjunctions, after demonstrative verbs, after certain numerals, after the predicative particle yr, and after Subjects (or, alternatively, on Objects or under c-command — see Harlow 1989, Borsley and Tallerman, in press).

In addition to the contexts listed in Table 1, SM plays a key role in the marking of grammatical gender, as follows:

1. Feminine singular nouns undergo SM after the definite article, yr(ð) and the numeral un ‘one’, e.g. y gath /sɡað/ (<cath /kaθ/ ‘the cat’; un gath /sɡaθ/ (<cath) ‘one cat’).

2. Adjectives undergo SM after feminine singular nouns, e.g. cath fawr /kaθ vaur/ (<mawr) ‘big cat’.

Note that SM occurs only when the feminine noun is singular. If a feminine noun occurs in the plural, neither it nor its modifying adjective undergoes soft mutation. This marking of feminine nouns in the singular contrasts with masculine nouns and adjectives following masculine nouns, which do not undergo Soft Mutation, e.g. y ci ‘the dog’; ci mawr ‘big dog’.

This use of SM to identify feminine nouns in ‘local’ contexts (i.e. on the noun and its immediate modifiers) contrasts sharply with the use of SM in ‘distant’ contexts for gender, with the possessive adjective ei (3rd person singular). Here SM signals a masculine antecedent, e.g. ei gath (/sɡað/) ‘his cat’, ei gi (/ɡi/) (<ci /ki/) ‘his dog’. If the antecedent is a feminine noun, the modified noun undergoes Aspirate Mutation, whereby /p, t, k/ become /f, θ, x/ (ei cɑth (/kaθ/) (<kɑθ/) ‘her cat’, ei chi (/xi/) (<kɪ/) ‘her dog’.

VIRGINIA C. MUELLER GATHERCOLE, ENLLI MÔN THOMAS & NADINE LAPORTE
### Table 1: Sample of environments that trigger Soft Mutation

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Examples</th>
</tr>
</thead>
</table>
| **Certain prepositions** | Mae Mair ar gadair (< cadair, F)  
'Mary is on (a) chair' |
| **ar** — 'on'      | Dw i'm mynd i llwyddo (< llwyddo)  
'I'm going to succeed' |
| **i** — 'to'       |                                                                           |
| **Certain possessives** | Lle mac dy bèl? (< pêl, F)  
'Where's your ball?' |
| **dy** — 'your'    | Saf ar dy draed (< troed, F, M)  
'Stand on your feet' |
| **Certain conjunctions** | Tatws neu foron? (< moron, F)  
'Potatoes or carrots?' |
| **neu** — 'or'     | Pan ddigwyddith (< digwyddith)  
'When it happens ...' |
| **Pan** — 'when'   |                                                                           |
| **Demonstrative verbs** | Dyna wel ( < gwely, M)  
'There's (a) bed' |
| **Dyna** — 'there is' | Dyma row (< rhaw, F)  
'Here's (a) shovel' |
| **Dyma** — 'here is' |                                                                           |
| **Certain numerals** | dau flodyn (< blodyn, M)  
'two flower(s)' |
| **Dau** — 'two' (< M) | dwy fasged (< basged, F)  
'two basket(s)' |
| **Dwy** — 'two' (< F) |                                                                           |
| **Predicate particle** | Roedd y cafu 'n wag (< gwag)  
'the cafe was empty' |
| **Yn** — 'two' (< M) |                                                                           |
| **Subject**        | Mi welod cath gi (< ci, M)  
'A cat saw a dog' |
| **(or, on the Object)** | Pwy welod mam?  
'Who did Mom see?' |
|                    | Pwy welodd fam? (< mam, F)  
'Who saw a Mom?' |

The resulting system is thus quite opaque, for the following reasons: (1) There is no one-to-one correspondence between form and function (e.g., we can't say that SM means 'feminine gender', or even 'some type of gender'). (2) There are gaps in the system (e.g., plural feminine nouns do not undergo or trigger SM). (3) There are contradictory triggers (both feminine and masculine gender can trigger SM). (4) There is a great deal of variability in adherence to the mutations in adult speech (Thomas 1984, Ball 1988b, Watkins 1993). There is variability across the three types of mutation (Ball 1988a, 1988c; Watkins 1993); within mutation types, by lexical item (Thomas 1984, Ball 1988c: 77-78, Hatton 1988, Thomas 2001); and in phonology, whereby phonemes susceptible to a given type of mutation are not equally affected in every context (Watkins 1993, Thorne 1993). (5) Finally, there are cases in which there is no overt marker of gender — namely, words beginning with non-mutable sounds. Words like *fenestr* 'window' (feminine) will have no overt marking of gender, since /f/ (like the other non-mutable word-initial consonants, /b/, s, f, h, v, n, l/; see Awbery 1984) does not undergo mutation.

Research that has been conducted on the acquisition of mutations in Welsh and other Celtic languages provides some insight into their development. First, there is considerable evidence that the acquisition of the mutation system is quite protracted and is not completed until well after age 5, perhaps even after age 11 (Bellin 1984, Hatton 1988, Jones 1992). Some have argued, in fact, that the mutation system in relation to gender is dying (Jones 1998). Second, initial steps to acquisition may be quite piecemeal. For example, Ó Baoill (1992) and Stephens (1996) have suggested that children may begin initially with a single form per noun. Third, children perform distinctly on formal versus naturalistic tasks, perhaps better on the former (Bellin 1988), perhaps worse (Thomas 2001). Fourth, the amount and type of linguistic input affects acquisition. Children who hear Welsh at home perform differently from children who hear English at home. The difference in performance is not necessarily one way, however (Hatton 1988, Jones 1992). Finally, bilingualism may affect acquisition. Two studies mention that the words that were least often mutated by their bilingual subjects were cognates in the two languages the children knew (Bellin 1988 for Welsh-English, Stephens 1996 for Breton-French). The following experiments were conducted to provide further information on these issues.

### Experiment 1: gender production

The purpose of Experiment 1 was to elicit children's productive command
of gender marking in a controlled experiment in which particular gendered constructs were elicited. The goal was to determine the following:

1. To what extent, and at what phase in development, do Welsh-speaking children rely on specific cues for gender categorization that are available in the input?

2. Is one type of contextual information more useful as a cue than another?

3. To what extent is children's knowledge of gender productive?

Method

Six tasks were designed to elicit the production of nouns and adjectives in gender-relevant contexts. Specifically, the tasks were designed to test knowledge of the marking of gender on the noun after the definite article and on the adjective following a noun.

Non-linguistic stimuli

Pictures were drawn to accompany the linguistic stimuli. For each noun tested, one or two pictures were shown to children. In some cases, one picture was shown to accompany both the stimulus given by the experimenter and the structure elicited from the child; in other cases, two pictures were used, one to accompany the stimulus given by the experimenter and the other to accompany the structure elicited from the child. The pictures for the given task were either presented together as a 'story' or shown as a series of 'cards' that the child could then 'post' in a box after responding.

Linguistic stimuli

Six tasks were designed, each of which provided the child with one linguistic structure with a given noun and asked the child to produce another structure with the same noun. Half the nouns included in the tasks were real nouns, and half were nonsense nouns, as follows. One hundred forty-four real words were selected for use across the six tasks. These included 6 masculine and 6 feminine nouns beginning with each of the consonants /p, t, k, b, d, g, v, r, m, n, f/, as well as 6 masculine and 6 feminine nouns that began with vowels. In addition, 144 nonsense nouns were constructed so that for each of the above consonant phonemes there were 12 nonsense words beginning with that sound; 12 nonsense words beginning with vowels were also created. This gave a grand total of 288 nouns. It should be noted that all of these nouns referred to inanimate objects.

Within a given task, each child heard 48 of these nouns (one masculine real for each of the word-initial phonemes, one feminine real for each of the phonemes, and two nonsense nouns for each of the phonemes), with the 288 words distributed across the six tasks for a given child. So that the distribution of the nouns across the six tasks was balanced, six versions of the entire test were drawn up. Version A had nouns 1-48 used in task 1, nouns 49-96 in task 2, nouns 97-144 in task 3, nouns 145-192 in task 4, nouns 193-240 in task 5, and nouns 241-288 in task 6; version B had nouns 49-96 in task 1, nouns 97-144 in task 2, and so forth; version C had nouns 97-144 in task 1, nouns 145-192 in task 2; and so on. Two children in each age group heard a given version of the test. Thus, each noun was heard in each task by two children in each age group, and each noun was used in every task an equivalent number of times across children.

The structures given by the experimenter and elicited back from the child for each of the six tasks are shown in Table 2 below. These were designed so that the structure provided by the experimenter contained no cue to the gender of the noun (tasks 1 and 2), one cue to the gender of the noun (presence or absence of SM on the noun after the definite article or on the adjective after a noun in tasks 3, 4, and 5), or two cues to the gender of the noun (presence or absence of SM on both the noun and a modifying adjective in task 6). Sample stimuli for each task are shown in Table 3.

Procedure

Each child was seen individually. For each task, the child was asked if he or she would like to play a game. The experimenter gave the child a sentence with one construction, and the child was asked to produce the target construction. For Task 1, for example, the child was asked if he or she would like to help tell a story about Mair, who had lost some things but then found them. The experimenter showed the child the first picture accompanied by the words Roedd Mair wedi colli Y 'Mair had lost an X'. The child's task was to produce Ond nath Mair wyd y Y 'but Mair saw the X'. Several practice items were used to ensure that the child understood what type of response s/he was expected to give.
**Table 2: Production of Gender, Elicitation Tasks (Experiment 1)**

<table>
<thead>
<tr>
<th>Task</th>
<th>Child was given</th>
<th>Form elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1: N → y N</td>
<td>noun</td>
<td>y + noun</td>
</tr>
<tr>
<td>Task 2: N → NA</td>
<td>noun</td>
<td>noun + adjective</td>
</tr>
<tr>
<td>Task 3: y N → NA</td>
<td>y + noun</td>
<td>noun + adjective</td>
</tr>
<tr>
<td>Task 4: NA → y N</td>
<td>noun + adjective</td>
<td>y + noun</td>
</tr>
<tr>
<td>Task 5: y N → N</td>
<td>y + noun</td>
<td>noun (without article)</td>
</tr>
<tr>
<td>Task 6: y NA → N</td>
<td>y + noun + adjective</td>
<td>noun (without article)</td>
</tr>
</tbody>
</table>

*ν = definite article; N = noun; A = adjective*

**Table 3: Sample stimuli (Experiment 1)**

<table>
<thead>
<tr>
<th>Task</th>
<th>Experimenter</th>
<th>Child</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>Roedd Mair wedi colli pêl 'Mair had lost a ball'</td>
<td>Ond nath Mair weld y bêl 'but Mair saw the ball'</td>
</tr>
<tr>
<td>Task 2</td>
<td>Roedd Mair wedi cael pêl 'Mair had got a ball'</td>
<td>Roedd Mair efo pêl goch rwan 'Mair had a red ball now'</td>
</tr>
<tr>
<td>Task 3</td>
<td>Ar y cerdyn yma fedri di weld y bêl 'on this card you can see the ball'</td>
<td>Ac ar y cerdyn yma fedrwych chi weld y bêl goch 'and on this card you can see the red ball'</td>
</tr>
<tr>
<td>Task 4</td>
<td>Roedd Mair wedi colli pêl goch 'Mair had lost a red ball'</td>
<td>Ond nath hi weld y bêl ddu ar ei gwely 'but she saw the black ball on her bed'</td>
</tr>
<tr>
<td>Task 5</td>
<td>Yn y llun yma fedri di weld y bêl 'in this picture you can see the ball'</td>
<td>Dyma lun pêl 'this is a picture of a ball'</td>
</tr>
<tr>
<td>Task 6</td>
<td>Yn y llun yma fedri di weld y bêl goch 'in this picture you can see the red ball'</td>
<td>Dyma lun pêl 'here's a picture of a ball'</td>
</tr>
</tbody>
</table>

**Predictions**

The predictions of this experiment were the following:

1. If children are using cues to gender status to inform themselves about the gender of a noun, they should perform better on tasks 3-6 (with cues) than on tasks 1 and 2 (without cues).

2. If children are using cues to gender status, they should perform better on tasks with multiple cues than on tasks with only one cue, i.e. better on task 6 (two cues) than on task 5 (one cue).

3. If gender is a property of nouns, then children should perform better at an early age on tasks eliciting noun forms than on tasks eliciting adjective forms, i.e. better on tasks 1, 4, 5, 6 (nouns) than on tasks 2 and 3 (adjectives).

4. Children should perform better on items involving real nouns than on those involving nonsense nouns.

5. However, if children are gradually developing an abstract, rule-based system, as they get older, performance on nonsense nouns should improve.

**Participants**

Forty-eight children took part in this study. The children attended primary schools either on Anglesey or in the Bangor and Ogwen area. All of them spoke Welsh as their first language and lived in homes where 80% to 100% of the input was in Welsh. (In fact, in all but two cases, parents reported 100% use of Welsh in the home.) The children were divided into four age groups, with 12 children in each group: Group 1 (4 1/2 year olds) were between 4;0 and 5;5, mean age = 4;9; Group 2 (6-year-olds) were between 5;7 and 6;7, mean age = 6;2; Group 3 (7 1/2-year-olds) were from 6;9 to 7;9, mean age = 7;2; and Group 4 (9-year-olds) were between 8;3 and 9;7, mean age = 8;9. Twenty-seven children were female and twenty-one were male.

**Results**

Each child was given a score in each cell for the proportion of times (or of that child's number of attempts) that he or she produced the appropriate form of noun or adjective. Since the data were in proportions, arcsine transformations were applied, and a repeated measures ANOVA was conducted on those scores. Task (tasks 1, 2, 3, 4, 5, and 6), gender (masculine feminine), noun type (real, nonsense), and age (4 1/2-year-olds, 6-year-olds, 7 1/2-year-olds, 9-year-olds) were the factors.
olds, 7 1/2-year-olds, and 9-year-olds) were treated as independent variables.

The analysis revealed significant main effects, first, of gender ($F (1, 33) = 874.41, p < .001$) and noun type ($F (1, 33) = 22.87, p < .001$). Performance was generally good on masculine forms ($M = .945, SD = .143$) and poor on feminine forms ($M = .202, SD = .262$); performance was better on real nouns ($M = .620, SD = .402$) than on nonsense nouns ($M = .527, SD = .448$).

There was also a significant main effect of task ($F (5, 165) = 9.27, p < .001$). Performance by task is shown in Figure 1 below. Planned mean comparisons revealed significantly better performance on tasks 5 and 6, on the one hand, than on tasks 1, 2, and 3 on the other, on task 4 than on task 2; and on task 6 than on task 4 (all $F$s (1, 165) $\geq 7.8$, all $p < .01$). These results indicate that children were better overall at giving the basic form of nouns when given gender-marked contexts (tasks 5 and 6) than they were at producing the appropriate form of nouns after the definite article when not given a cue (task 1). They were also better at giving the basic form of nouns given gender-marked contexts (tasks 5 and 6) than they were at producing the appropriate form of adjectives (tasks 2 and 3). The children were also better at producing the appropriate form of nouns after the definite article in task 4 than they were at producing the appropriate form of adjectives in task 2. Finally, the children were also better at giving the basic form of nouns from gender-marked contexts (task 6) than they were at producing the appropriate form of nouns after the definite article, even when a cue was given (task 4).

These main effects were modified by several significant interactions: task x gender ($F (5, 165) = 9.71, p < .001$); noun type x gender ($F (1, 33) = 15.51, p < .001$); task x noun type ($F (5, 165) = 2.50, p < .05$); gender x age ($F (3, 33) = 3.75, p < .05$); task x gender x age ($F (15, 165) = 2.53, p < .01$); and task x noun type x gender ($F (5, 165) = 4.32, p < .01$).

Performance by task and gender is shown in Figure 2. Follow-up analyses revealed that with masculine forms children performed better on tasks 4, 5, and 6 than on tasks 2 and 3, and with feminine forms, better on tasks 2 and 3 than on task 1, and on task 3 better than on task 4 (all $F$s (1, 165) $> 5.6$, all $p < .05$). That is, with masculine forms, children were better at giving the basic forms of nouns, either with or without the article (tasks 4, 5, and 6) than they were at producing the appropriate (basic) forms of adjectives (tasks 2 and 3). With feminine forms, in contrast, children were better at producing the mutated forms of adjectives after nouns (tasks 2 and 3) than they were at producing the mutated forms of nouns after the definite article given no cue (task 1). They were also better at producing

the SM form of adjectives after feminine nouns given a cue (task 3) than they were at producing the SM form of feminine nouns after the article also when given a cue (task 4).

Figure 1: Performance by task, Experiment 1

![Figure 1: Performance by task, Experiment 1](image1)

Figure 2: Performance by task x gender, Experiment 1

![Figure 2: Performance by task x gender, Experiment 1](image2)

The interaction of noun type x gender, shown in Figure 3, indicated that
the better performance on real than on nonsense nouns occurred only with
feminine forms \( F(1,33) = 56.9, p < .001 \).

**Figure 3: Performance by noun type x gender, Experiment 1**

The interaction of task x noun type is shown in Figure 4 below. Follow-
up analyses revealed that for real nouns, children performed better on
tasks 5 and 6 than on tasks 1, 2, 3, and 4, and for nonsense nouns, they
performed better on tasks 5 and 6 than on task 2, and better on task 3 than
on tasks 4 and 6 (all \( Fs(1, 165) > 6.0 \), all \( p < .05 \). That is, for real nouns,
children performed best when giving the basic forms of nouns (tasks 5 and
6), and were better at this than they were either at producing the appro-
priate forms of nouns after the definite article (tasks 1 and 4) or at giving
the appropriate forms of adjectives (tasks 2 and 3). For nonsense forms, they
were better at giving the basic forms of nouns (tasks 5 and 6) than they
were at producing the appropriate forms of adjectives, given no cue (task
2), and they were better at producing adjectives, given a cue to gender
(task 3), than at producing noun forms, either after the definite article
given a cue (task 4) or in basic form after gender-marked cues (task 6).

Performance by gender and age is shown in Figure 5. Follow-up analy-
ses revealed that there were no significant differences across ages for per-
formance on feminine forms, but for masculine forms, 7 1/2-year-olds per-
formed better than 9-year-olds (Student-Newman-Keuls analysis, \( p < .05 \).
The three-way interaction of task x gender x age, shown in Figures 6 and
7, further revealed that with masculine forms (Figure 6), 6- and 9-year-
olds performed better on tasks 1, 4, 5, and 6 than on task 2, and 9-year-
olds also did better on tasks 1, 4, 5, and 6 than on task 3 (all \( Fs(1, 165) >
6.4, all \( p < .05 \)). Thus, on masculine forms, 6- and 9-year-olds were better
at producing (the basic forms of) nouns than at producing (the basic forms
of) adjectives (task 2). For 9-year-olds, this was the case even when a gen-
der cue was present (task 3). On feminine forms, 9-year-olds performed
better on the production of (mutated forms of) adjectives when given a cue
(task 3) than on the production of the noun form (tasks 1, 4, 5, and 6).
Figure 6: Performance by task x age x gender, Experiment 1 (masculine forms)

Figure 7: Performance by task x age x gender, Experiment 1 (feminine forms)
Finally, the interaction of task x noun type x gender is shown in Figure 8. Follow-up analyses revealed that, for real masculine nouns, performance was better on task 6 than on tasks 2 and 3, and on 1 than on 3 (Student-Newman-Keuls analysis, $p < .05$). Thus children generally performed better at giving the (basic) forms of nouns, with or without the article (tasks 6 and 1) than at producing the (basic) forms of adjectives (tasks 2 and 3). For real feminine nouns, children were better at giving the basic form of nouns after having cues (task 6) than they were at producing the SM form of feminine nouns after the article, after having no cue (task 1) (Student-Newman-Keuls analysis, $p < .05$).

For nonsense masculine forms, performance was better on tasks 1, 4, 5, and 6 than on tasks 2 and 3 (Student-Newman-Keuls, $p < .05$), i.e., the worst performance was at producing (basic forms of) adjectives after the nouns. For nonsense feminine forms, performance was better on task 3 than on tasks 1 and 4 (Student-Newman-Keuls, $p < .05$), that is, children were better at producing the (SM form of) adjectives when given a cue (task 3) than they were at producing the (SM form of) nouns after the definite article (tasks 1 and 4).

**Summary and discussion: Experiment 1**

Overall, the children performed better on masculine forms than on feminine forms. They were also better when real feminine nouns occurred than when nonsense feminine nouns were used. Across tasks, children were better at giving the basic forms of nouns without articles than they were at producing the appropriate form of nouns after the definite article or of adjectives after nouns.

Performance was different, however, with masculine versus feminine forms: with masculine forms, children were worse at providing appropriate (non-mutated) forms of adjectives than of noun forms. With feminine forms, in contrast, children performed better at giving appropriate (mutated) forms of adjectives than appropriate (mutated) forms of nouns after the definite article. Thus, children appeared to be overextending the use of mutated adjectives to masculine contexts, and to be underextending the use of SM for nouns in feminine contexts. These effects were most apparent among the 6- and 9-year-olds.

The results of this experiment shed light on the predictions made at the outset, as follows:

1. If children are using cues to gender status to inform themselves about the gender of a noun, they should perform better on tasks 3-6.
than on tasks 1 and 2. This prediction was not upheld by the data. While there were differences across tasks, they follow a more complex pattern, most often involving differential performance on nouns versus adjectives or on mutated versus non-mutated forms.

(2) If children are using cues to gender status, they should perform better on tasks with multiple cues than on tasks with only one cue, i.e. better on task 6 than on task 5. This prediction was similarly not upheld by the data.

(3) If gender is a property of nouns, then children should perform better at an early age on tasks involving noun form than on tasks involving adjective form, i.e. better on tasks 1, 4, 5, and 6 than on tasks 2 and 3. This prediction also did not in general hold. With masculine forms, performance was better on nouns than on adjectives, but with feminine forms, performance was better on adjectives than on nouns. This appears related to overuse of mutated forms of adjectives, even in masculine constructs, and underuse of mutated forms of nouns in feminine constructs with the definite article.

(4) Children should perform better on items involving real nouns than on those involving nonsense nouns. This prediction was upheld by the data.

(5) However, if children are gradually developing an abstract, rule-based system, as they get older, performance on nonsense nouns should improve. This prediction was not upheld by the data. Performance on nonsense nouns did not improve with age.

The results of this experiment suggest overall that perhaps children are not developing a system per se for gender: they do poorly on nonsense nouns, and a greater number of cues to gender status do not appear to boost performance. Instead, children's knowledge appears to be more piecemeal, with no strict abstract rules that govern the production of all forms. That more piecemeal knowledge may include the following: (a) For masculine nouns, children use a single (basic) form in these constructs. (b) Children seem to know that adjectives often occur in mutated forms in these constructs. This knowledge appears to grow with age, so that there is a dramatic increase at age 9 in the correct use of mutated adjectives with feminine nouns (task 3). This growing knowledge would also account for the overuse of mutated adjectives in masculine constructs, especially in two of the older groups. (Such overuse is also in evidence among adults; cf. Thomas 2001.) (c) Children appear less knowledgeable that feminine nouns mutate after definite articles. Performance on feminines in this context did not improve with age. We shall return to these points in the final discussion section.

Experiment 2: gender comprehension

The purpose of Experiment 2 was to elicit children's interpretations of long-distance gender linkages between nouns and coreferential pronouns and possessive forms. The questions addressed were the following:

(1) To what extent are children's interpretations of gendered pronouns (hi, feminine, and o, masculine) and of the third person singular possessives (ei) dependent on the grammatical gender of an antecedent, and is there a change with age?

(2) Is there a difference in children's reliance on grammatical gender in relation to pronominal forms (which do not involve mutations) vs. the possessive form (which does)?

(3) Is there a difference in children's abilities with the two types of forms in relation to human, animal, and inanimate antecedents?

(4) Finally, is there a difference in children's abilities according to the language used in the home, whether Welsh only is spoken, both Welsh and English are spoken, or English only is spoken?

The answers to these questions will help to illuminate, first, the degree to which children gain a productive command of these forms. Secondly, the data will provide information on the relative status of pronouns and possessives as involving natural gender versus grammatical gender. If children interpret these forms as referring to natural gender, i.e., as corresponding to the real-world sex of referents, they should perform better on human antecedents than inanimate antecedents (with animal antecedents somewhere between these). In contrast, if they interpret the forms as involving grammatical gender, performance should be equivalent for human, animal, and inanimate antecedents. Further, the data will provide some evidence.

---

3. Experiment 2 is part of a larger ongoing study into the acquisition of Welsh by Welsh-dominant, balanced, and English-dominant bilinguals in North Wales.
information on the role that the language of the home might play in the acquisition of these forms.

Method

The general method involved a forced-choice task between two pictures, to find the one that corresponded to the sentence uttered by the experimenter. Within each trial, participants were shown an initial picture that corresponded to an initial sentence involving two nouns, one masculine and one feminine, and then two pictures from which they had to choose one to match a second sentence that included a masculine or feminine pronoun or possessive form.

Non-linguistic stimuli

Sets of pictures were drawn to accompany each trial. In each trial set, an initial picture showed two referents corresponding to two different nouns, one masculine and one feminine, e.g. a dish (dysgl, F) and a flower (bloody, M). This first picture was on one side of a card and was to be shown while the experimenter uttered a first sentence with the two relevant nouns. The trial set also included a pair of contrasting pictures, depicted on the other side of the card. These were shown when a second sentence, containing either a M or F pronoun or possessive form, was uttered. One picture corresponded to an interpretation of the pronoun or possessive as coreferential with the feminine noun in the first sentence, and one corresponded to an interpretation of it as coreferential with the masculine noun. That is, in one of the two pictures, the referent for the masculine noun was depicted with the relevant property (e.g. being broken) expressed in the second sentence; in the other, the referent for the feminine noun was depicted with the relevant property expressed in the second sentence. There were 36 such stimulus sets, corresponding to the 36 linguistic stimuli.

Linguistic Stimuli

Linguistic stimuli were prepared, with each containing two sentences. In the first sentence, a masculine and a feminine noun occurred; in the second, either a masculine or a feminine pronoun or possessive form occurred. The two nouns that occurred in the first sentence fell into three referent types: two nouns referring to humans, two nouns referring to animals, or two nouns referring to inanimate objects. Examples of test sentences are shown in Table 4, which differentiates between sentences used to test the pronoun interpretations and those used to test the interpretation of ei.

Table 4: Sample sentences used as stimuli in Experiment 2

<table>
<thead>
<tr>
<th>NOUNS FOR HUMANS</th>
<th>NOUNS FOR ANIMALS</th>
<th>NOUNS FOR INANIMATES</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Roedd gwr blin (M) a'r gyfnith ddireidus (F) ar ben y sfeid. Aeth o/hi i lawr y sfeid gyntaf.</td>
<td>(2) Aeth y pengwin cysglod (M) a'r gath flinedig (F) i'r gwelyau. Aeth o/hi i gysgu gyntaf.</td>
<td>(3) Roedd y ddysgllas (F) a'r blodyn piws (M) ar y bwrrd. Ond nath o/hi dorri.</td>
</tr>
<tr>
<td>'The grumpy man and the mischievous cousin (F) were on top of the slide. He/she went down the slide first.'</td>
<td>'The sleepy penguin and the tired cat went to their beds. He/she went to sleep first.'</td>
<td>'The blue dish and the purple flower were on the table. But it (M/F) broke'.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NOUNS FOR HUMANS</th>
<th>NOUNS FOR ANIMALS</th>
<th>NOUNS FOR INANIMATES</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4) Dyma'r gwfir pwysig (M) a dyma'r ddyonas dda (F). Mae yna het ar ei phen [ei + AM, therefore F]</td>
<td>(5) Dyma'r gath dew (F) a dyma'r draenog pigog (M). Ond mae 'na bili pala ar ei thrwyn. [ei + AM, therefore F]</td>
<td>(6) Dyma'r wasgod ddu (F) a dyma'r crys pinc (M). Ond ma'i phoed yn las. [ei + AM, therefore F]</td>
</tr>
<tr>
<td>'Here's the important man and here's the good woman. There's a hat on her head.'</td>
<td>'Here's the fat cat and there's the prickly hedgehog. But there's a butterfly on its (F) nose.'</td>
<td>'Here's the black waistcoat/vest and here's the pink shirt. But its (F) pocket is blue.'</td>
</tr>
</tbody>
</table>

The initial sentence in every case was designed so that the nouns were introduced with the definite article y(r) and a modifying adjective. Furthermore, these nouns and the adjectives modifying them all began
with one of the following sounds susceptible to Soft Mutation: /p, t, k, b, d, g/. This was done to ensure that in every case, the first sentence would contain clear clues to each noun's gender, since feminine nouns after *y(r)* undergo SM and trigger mutation on the following adjective, while masculine nouns do not.

In the case of the stimuli testing the interpretation of *ei*, the noun modified by *ei* in the second sentence always began with /p/, /t/, or /kl/. These are the only three sounds susceptible to both SM and Aspirate Mutation; hence, the construct in every case was clearly marked for masculine (*ei* + Soft Mutation) or feminine (*ei* + Aspirate Mutation) gender.

Each participant had a total of 36 stimulus trials. These trials consisted of 12 sentences with nouns for humans, 12 with nouns for animals, and 12 with nouns for inanimate objects. Within each group of 12, six tested interpretations of pronouns, and six tested interpretations of *ei*. Within each of the six trials on pronouns, half tested the interpretation of *hi*, the feminine pronoun, and half *o*, the masculine pronoun. Similarly, within each of the six trials on *ei*, half tested the feminine *ei* (i.e. with Aspirate Mutation) and half tested the masculine *ei* (i.e. with Soft Mutation). Furthermore, within each of the six trials in each group, three trials presented the masculine noun before the feminine noun in the first sentence (as in sentence 1 in Table 4), while three presented the feminine noun first (as in sentence 3 in Table 4).

Four versions of the 36 stimuli were drawn up, to ensure balancing of stimuli. The four versions (1a, 1b, 2a, 2b) were constructed as follows. First, sets I and II were identical, except that the order in which the masculine and feminine nouns occurred in the first sentence for each trial of set I was reversed for the comparable sentence of set II (e.g., set I had stimulus 1 in Table 4, set II had its reverse, with the feminine noun before the masculine noun). Second, the (a) and (b) versions of each set were identical except that in the second sentence of a stimulus, one version had the masculine form and the other had the feminine form. Thus, for example, set 1a had stimulus 1 in Table 4 with the masculine form, *o*, while set 1b had the same stimulus, but with the feminine form, *hi*. Finally, the order in which the stimuli were presented was determined as follows. Each of the four versions of stimuli was randomized, and then half of the children received the stimuli in top-to-bottom order, and half received them in bottom-to-top order.

**Procedure**

Each child was seen individually and sat next to the experimenter. Children were asked, in Welsh, to play a game. The experimenter showed a picture and uttered a sentence corresponding to it. She then turned over the card, to show the two choice pictures, and uttered the second sentence. The child was instructed to point to the picture that the experimenter was 'talking about'. Two practice items that were not relevant to the task were used; these were followed immediately by the trial items.

**Participants**

The data presented here come from a preliminary set of 101 participants. These fell into three age groups: 26 'five-year-olds' (mean: 5;7, range: 4;8-6;5), 37 'seven-year-olds' (mean: 7;10, range: 6;9-8;10), and 38 'nine-year-olds' (mean: 9;11, range: 8;11-11;1). Fifty-five were female; 46 were male. The children came from one of three types of homes: homes in which Welsh was spoken over 80% of the time since the child was born (N=60), in which English was spoken over 80% of the time since the child was born (N=15), or in which both Welsh and English were spoken 40% to 60% of the time since the child was born (N=26). Approximately half the children were classified as high SES (N=50) and half as lower SES (N=51), based on the parents' educational levels and professions.

**Results**

An ANOVA was conducted in which age (5, 7, 9); home language (OWH [only Welsh at home], OEH [only English at home], and WEH [Welsh and English at home]); and SES (high, low) were treated as between-subjects variables. Linguistic form (pronoun, possessive), animacy (human, animal, inanimate), and gender (feminine, masculine) were within-subjects variables. Results revealed the following significant effects.

First, there were main effects of age (*F*(2, 85)=9.54, *p*<.0002), home language (*F*(2, 85)=4.39, *p*<.02), linguistic form (*F*(1, 85)=6.06, *p*<.02), and animacy (*F*(2, 170)=60.78, *p*<.0001). The effect of age was due to the fact that the responses of the children in the three age groups all differed from one another (Student-Newman-Keuls analysis, *p*<.05), with 60.9% accuracy at age 5, 65.7% at age 7, and 70.1% at age 9. The effect of home language revealed significant differences in performance by children coming from OEH homes (at 61.5% accuracy) versus OWH homes (68.2% accuracy), Student-Newman-Keuls analysis, *p*<.05. (Those coming from WEH homes fell between these two, at 63.9% accuracy, which was not significantly different from either of the other groups.) The main effect of linguistic form revealed that there was better performance overall on the
pronouns (67% correct) than on the possessive (65.2%). Finally, the main effect of animacy was due to better performance on sentences involving human referents (87.2% correct) than on those involving either animal referents (57.1% correct) or inanimate referents (54.0% correct) — human vs. animal: \( F(1, 170)=81.5, p<.0001 \), human vs. inanimate: \( F(1, 170)=99.9, p<.0001 \). There were no other significant main effects.

There were also several significant two- and three-way interactions: linguistic form x animacy, \( F(2, 170)=17.53, p<.0001 \); animacy x age, \( F(4, 170)=3.87, p<.005 \); gender x animacy, \( F(2, 170)=4.67, p<.02 \); linguistic form x home language, \( F(2, 85)=3.21, p<.05 \); linguistic form x gender x home language, \( F(2, 85)=3.32, p<.05 \); linguistic form x gender x animacy, \( F(2, 170)=7.34, p<.0009 \); and linguistic form x animacy x home language, \( F(4, 170)=5.45, p<.0004 \).

Figure 9 shows performance according to linguistic form and animacy. Follow-up analyses revealed significant differences in performance on pronouns when they referred to the three types of antecedents (all \( Fs \geq 8.9, p<.004 \)), and significant differences in performance on the possessive when it referred to human antecedents versus animal or inanimate antecedents (both \( Fs \geq 18.4, p<.0001 \)).

**Figure 9: Performance by linguistic form x animacy, Experiment 2**

Performance by animacy and age is shown in Figure 10. There was clear development by age in relation to human referents, a slight improvement with age in relation to animal referents, and no change with age in relation to inanimate referents. It should be noted that the improvement with age in relation to humans and animals occurred both with the pronouns and with the possessive (humans: pronouns: 80%, 90%, and 100% at the three ages, respectively; possessives: 71%, 81%, and 90%, respectively; animals: pronouns: 52%, 55%, 61%, respectively; possessives: 53%, 57%, 62%, respectively).

**Figure 10: Performance by animacy x age, Experiment 2**

Performance by gender and animacy is shown in Figure 11. Follow-up analyses revealed significant differences on the three types of referents with feminine forms (all comparisons, \( F(1, 170)=8.1, p<.005 \)) and between forms with human referents versus animal or inanimate referents with masculine forms (both \( Fs \geq 37.2, p<.0001 \)).

**Figure 11: Performance by gender x animacy, Experiment 2**

Performance by linguistic form and home language is shown in Figu
12. While the OWH group performed similarly on the two constructs, the OEH and WEH groups performed better on the pronouns than on the possessive.

*Figure 12: Performance by linguistic form and home language, Experiment 2*

The three-way interactions of linguistic form and home language with animacy and with gender, shown in Figures 13 and 14, illuminate the source of these interactions. Figure 13 shows performance by animacy and linguistic form, broken down by home language. Follow-up analyses revealed that there was no difference across home language groups in performance on pronouns by animacy, but there was in performance on the possessive by animacy, $F(4, 196) = 4.4, p < .002$. Specifically, the OWH group performed better on the possessive in relation to humans and animals than the other two groups.

The significant interaction of linguistic form x gender x home language (Figure 14) is due to the fact that within the OWH group, there was a significant difference in performance on the possessive in relation to feminine versus masculine antecedents, $F(1, 59) = 28.2, p < .0001$. (The other two home language groups showed no significant difference there.) Examination of the data reveals that while all three groups performed similarly on the possessive with masculine antecedents, the OWH group performed better on the possessive with feminine antecedents, especially antecedents that referred to humans and animals. (Performance on the possessive with feminine human antecedents was OWH: 91%, OEH: 71%, and WEH: 80%. Performance with feminine animal antecedents was OWH: 78%, OEH: 51%, and WEH: 59%. With feminine inanimate antecedents performance on the possessive was uniformly low — OWH: 54%.

*Figure 13: Performance by animacy x linguistic form x home language, Experiment 2*

Finally, the performance by linguistic form, gender, and animacy is...
shown in Figure 15. Follow-up analyses revealed that there was no significant difference in performance by gender on the pronouns, but there was a significant difference by gender in interpretation of the possessive in relation to animals and inanimates ($F$s (1, 170) ≥ 9.8, $p < .002$). Specifically, performance was better on feminine than masculine possessives in relation to animals ($F(1,170)=16.9, p<.0001$), but better on masculine than feminine possessives in relation to inanimates ($F(1,170)=9.8, p<.003$).

Figure 15: Performance by linguistic form x gender x animacy, Experiment 2

Summary and discussion, Experiment 2

These data reveal the following with regard to children's interpretations of long-distance constructs involving pronouns and the possessive ei. First, performance was best in relation to human antecedents and worst in relation to inanimate objects. Performance on animals was sometimes between these two, sometimes equivalent to inanimates. Second, there was a general progression with age, especially with regard to human, but also to animal, antecedents. That progression occurred in relation both to pronouns and the possessive. Third, overall performance was better in relation to the pronouns than the possessive. Fourth, the OWH group performed better than the other two groups, especially with regard to feminine uses of the possessive in relation to humans and animals.

These results shed some light on the questions posed at the outset, as follows.

(1) To what extent are children's interpretations of gendered pronouns (hi [F] and o [M]) and of the third person singular possessives (ei) dependent on the grammatical gender of an antecedent, and is there a change with age? The data indicate that children do have the capacity to interpret gendered forms in relation to the gender of an antecedent at a relatively early age, but this ability does improve with age. However, the exact nature of this ability is contingent on the type of form (see point 2) and the nature of the referent (point 3).

(2) Is there a difference in children's reliance on grammatical gender in relation to pronominal forms (which do not involve mutations) versus the possessive form (which does)? Children were better at responding to the pronominal forms than to the possessive form. However, there was improvement on responses to the possessive form, especially for feminines, and especially in the OWH group.

(3) Is there a difference in children's abilities with the two types of forms in relation to human, animal, and inanimate antecedents? These results clearly indicate better performance in relation to human referents than in relation to animals and inanimates. This difference applies to both pronominal forms and possessive forms.

(4) Finally, is there a difference in children's abilities according to the language used in the home, whether Welsh only is spoken, both Welsh and English are spoken, or English only is spoken? The data show differences between the OWH group and the other two groups, with OEH children performing worst overall. These differences are most likely due to frequency of input (see Gathercole, in press a,b,c for a similar suggestion with regard to Spanish-English bilinguals in the Miami area). The primary advantage of the OWH group appears to lie in the developing ability with the possessive form in relation to feminine referents, especially humans and animals.

General Discussion

These two experiments together provide some insight into the process by which Welsh-speaking children break into the complex grammatical gender system. Experiment 1 tested children's productive capacities in relation to local gender marking with masculine and feminine nouns for inanimate objects. The data suggested that children of these ages have not (yet) developed an abstract, rule-based system, but, rather, have a constellation of
piecemeal knowledge concerning the use of mutation in relation to gender. The results showed (a) consistently high use across ages of the basic forms of nouns when those nouns were masculine, (b) some overuse of mutated forms of adjectives in masculine contexts, especially at older ages, and (c) underuse of mutated forms of feminine nouns in contexts that require SM, i.e. after the definite article.

One might wonder whether one could interpret these results as indicating that Welsh-speaking children simply use basic forms of nouns for both masculine and feminine nouns (hence, a high level of performance on masculines and low on feminines) and mutated forms of adjectives as the 'default' form for adjectives (hence, the overextension of mutated adjectives to masculine constructs). A close examination of the data suggests that this would not give an accurate picture, however. First, the high performance on masculine nouns and lower performance on feminine nouns are not inverses of each other. So, for example, on task 4, Figure 7, children performed at 98% accuracy on giving the (basic form of) real masculine nouns after the definite article, but gave the basic forms of real feminine nouns after the definite article only 78.5% of the time. Similarly, the use of mutated forms of adjectives with real masculine nouns in task 3 occurred only 9.1% of the time, compared with 36.4% of the time with real feminine nouns. (In fact, among the 9-year-olds, who showed the greatest tendency to overextend mutation to adjectives in the masculine context, mutation of adjectives with real masculine nouns occurred only 21% of the time, while with real feminine nouns, it occurred 60% of the time.)

The results of Experiment 2 give a similar picture of piecemeal learning, this time in relation to gender marking in distant constructs involving pronouns and possessive forms. The data suggest that Welsh-speaking children may break into this system through gender marking in relation to nouns for human referents, linking gender to real-world, natural gender. If the knowledge of gender marking remained limited to use for humans, then one might conclude that the grammatical gender system of Welsh is becoming a natural gender system. One clue that this may not be the case is the developing command of gender in the OWH group, the most advanced of the three groups, in relation to possessives as they apply to animal referents. The discrimination of the masculine and feminine uses of the possessive in relation to animals goes beyond natural gender, because it is not the natural gender of the animal that determines the gender of the noun, but the particular conventional assignment of the noun to one grammatical gender or the other. (E.g., *ci ‘dog’, *mochyn ‘pig’, and *gloyn byw ‘butterfly’ are masculine, but *cath ‘cat’, *dafad ‘sheep’, and *ylluan ‘owl’ are feminine.)

There are two possible reasons why the OWH children differ from the OEH and WEH children in this regard. One is that the OWH children may be ahead of the other two groups, in which case we would expect the other children will eventually catch up and also begin to treat gender as grammatical gender (see such a suggestion by Oller, in press regarding ‘catching up’ by certain Spanish-English bilingual groups). The other possibility is that the OWH children may be less influenced by English than those who come from OEH and WEH homes. Since English is a natural gender language, perhaps greater dominance from English in the bilingual’s system may lead a child to tend towards a more natural gender approach than greater dominance in Welsh would. That is, perhaps in the end, the OWH group will be the only ones whose command of gender in Welsh moves away from natural gender towards more grammatical gender. This intriguing possibility will have to be left for further research.

The ultimate conclusion to be drawn from the two experiments is that when a language presents the child with a very complex and opaque system for gender, the course of development is protracted and variable. Children acquire SM for gender in a very piecemeal fashion, and the child appears to draw on multiple sources to construct that system. The differences between the near-perfect use of basic forms for masculine nouns and more varied use of mutated and non-mutated forms for feminine nouns and for adjectives modifying both masculine and feminine nouns may have their source in the differences in the frequency and consistency with which these distinct forms are heard in the input to the child. Similarly, the possible movement from a more natural gender system for humans to a more grammatical gender system through the gradual extension of the feminine possessive form to nouns for animals may stem from the fact that the feminine possessive involves Aspirate Mutation, not Soft Mutation. While SM applies in a vastly varied set of distinct morphological and syntactic contexts, Aspirate Mutation is more confined to just a few constructs. Its application in the case of the feminine possessive form may thus provide a more reliable cue to gender than the application of, e.g., Soft Mutation in the case of the masculine possessive (cf., for example, MacWhinney, Bates, and Kliegl 1984; McDonald 1986; MacWhinney 1987; Sokolov 1989; and Bates and MacWhinney 1989). The feminine possessive, then, may provide the ‘window’ that children need to be able to look beyond the tangle of information in the input that Soft Mutation provides towards a clearer glimpse of a system governing grammatical gender in Welsh.
References


Oller, D. Kimbrough (ed.) in press. Language and Literacy in Bilingual Children.


