

Jennifer Smith, Mercedes Durham and Hazel Richards

The social and linguistic in the acquisition of sociolinguistic norms: Caregivers, children, and variation¹

Abstract: Despite the assumption in early studies that children are monostylistic until sometime around adolescence, a number of studies since then have demonstrated that adult-like patterns of variation may be acquired much earlier. How much earlier, however, is still subject to some debate. In this paper we contribute to this research through an analysis of a number of lexical, phonological and morphosyntactic variables across 29 caregiver/child pairs aged 2;10 to 4;2 in interaction with their primary caregivers. We first establish the patterns of use – both linguistic and social – in caregiver speech and then investigate whether these patterns of use are evident in the child speech. Our findings show that the acquisition of variation is highly variable dependent: some show age differentiation, others do not; some show acquisition of style shifting, others do not; some show correlations between caregiver input and child output, others do not. We interpret these findings in the light of community norms, social recognition and sociolinguistic value in the acquisition of variation at these early stages.

Keywords: acquisition, children, caregivers, variation, sociolinguistic

Jennifer Smith: English Language, School of Critical Studies, University of Glasgow,
12 University Gardens, Glasgow, G12 8QQ, United Kingdom
E-mail: jennifer.smith@glasgow.ac.uk

Mercedes Durham: Cardiff University

Hazel Richards: University of York

¹ We gratefully acknowledge the support of the Economic and Social Research Council for award no RES-000-22-0447 and the British Academy for award no. SG/45936. This research was also supported in part by the program *Apprentissages, connaissances et société* funded by the ANR (French national agency for research). We thank Paul Foulkes, Jean-Pierre Chevrot, and two anonymous reviewers for very helpful feedback on earlier versions of this manuscript. We would also like to thank Moira Smith, our invaluable on-site fieldworker, and the caregivers and children of Buckie who so warmly welcomed us into their *hooses*.

1 Introduction

Consider the exchanges in (1) between Luke, aged 3;5 and his mother Molly.

- (1) a. (Luke) *Are we gan to Isla's?* (Molly) *Uhuh.* (Luke) *Are we?* (Molly) *Later on, aye.* (Luke) *Say yes or no.* (Molly) *Aye ... yes.* (Luke) *No, say yes or no.* (Molly) *Yes.*
 b. (Luke) *Take it off!* (Molly) *No **don't** no! **Dinna** now! Just leave that. Just hold on. No **dinna**! You'll break it. **Don't** do that!* (Luke) *How?*
 c. (Molly) *Where's your shoes?* (Luke) *I dinna ken where they're at.* (Molly) *Your wellie boots **is** in the back of the car. You wanting them on again?* (Luke) *Aye.*

These interactions between mother and child raise a number of questions about acquisition of variable forms. At what age does Luke “know” that there are two ways of saying *yes* in his dialect: standard *yes* and the local form *aye*? When does Luke first “know” that there are two forms of the negative imperative – *dinna* and *don't*? When does Luke “know” that in his dialect he can say *The wellie boots is in the back of the car* but never *They is in the back of the car*? And when does Luke not only “know” about these patterns of variation, but also when will he start to use them in an adult-like structured way? In other words, when does Luke, and every other young child, start to acquire the highly complex patterns of variation, both linguistic and sociolinguistic, widely attested in adult speech?

Despite the assumption in early studies that children are monostylistic until sometime around adolescence (e.g., Labov 1970; Wolfram and Fasold 1974) a number of studies since then have demonstrated that adult-like patterns of variation may be acquired much earlier. How much earlier, however, is still subject to some debate. Some suggest that systematic patterns are acquired in the pre-adolescent years e.g., 10–12 (e.g., Reid 1978; Renn and Terry 2009; Romaine 1984; Chevrot et al. 2000), others in the first school years i.e., 6–8 years old (e.g., Labov 1989; Patterson 1992). Even more recent research provides some evidence that they are acquired even earlier, in tandem with the acquisition process more generally (e.g., Diaz-Campos 2005; Foulkes et al. 2005; Roberts 1994; Smith et al. 2007).

These differing results on when a child gains “sociolinguistic maturity” (Kerswill and Williams 2000: 105) may be in part due to the different linguistic variables under study. Kerswill (1996: 199) points out that “exactly when a child acquires a feature of his or her first dialect depends on the linguistic level [and] the complexity of the conditioning” of the variable in question. Chevrot et al. (2000: 296) suggest that in addition to linguistic considerations, the age at which

sociolinguistic patterns are acquired “depends on the perceptual salience of the variants in question [...] and their sociolinguistic value in a given community.” In other words, the complex linguistic and social correlates of a particular variable will have a significant effect on what is acquired when. However, “salience” and “sociolinguistic value” in a particular community might be “different for different social groups” (Kerswill and Williams 2002: 101). It might also be different for speakers within a group, because of competing appropriateness norms (e.g., Wolfram 1991). These “norms” are particularly relevant in the case of initial input in these early years as child directed speech (CDS) is demonstrated to be highly influential in the acquisition of variant forms (e.g., Kerswill and Williams 2000). At the same time CDS may differ from general community norms and differ according to the age of the child (e.g., Foulkes et al. 2005; Roberts 2002).

Our previous work on the acquisition of variation in pre-school children in interaction with their primary caregivers (Smith et al. 2007, 2009) indicates that sociolinguistic norms are evident from the earliest stages. We found that the specific linguistic variable under study, in tandem with the age of the child, had a significant impact on caregiver input and in turn, on how and when the acquisition of adult-like patterns of variation emerges in child output. However, we hypothesized that an overriding factor in explaining the contrasting patterns of acquisition rested on whether the variable being acquired had some “social recognition” (Labov 2001: 196) in the community in question.

In this paper we test this hypothesis further by bringing together results from previous analyses (Smith et al. 2007) with a number of new variables and a larger sample of speakers. In our previous work, we included 11 caregiver/child pairs. Here we include 29. In our previous research, we have investigated three variables. In this paper we investigate six, taken from different areas of the grammar. We first establish the patterns of use – both linguistic and social – in caregiver speech and then investigate whether these patterns of use are evident in the child speech. This will allow us to explore further a range of questions related to the acquisition of sociolinguistic norms in these early stages. We first review the data in question.

2 Caregivers and children in Buckie, Scotland

The data come from a small fishing town, Buckie (population circa 8000), which is situated on the north east coast of Scotland, 60 miles north-west of Aberdeen. The speech community forms a linguistically homogeneous group (e.g., Smith 2000a, 2000b, 2001) which provides a controlled environment for tracking the emergence of structured variation in child language. The caregiver/child corpus

Table 1: Caregiver/child corpus

| Child's name | Caregiver's name | Sex | Child's age at time of recording |
|--------------|------------------|-----|----------------------------------|
| Elizabeth | Martha | F | 2;10 |
| Jade | Avril | F | 2;10 |
| Ricky | Sheila | M | 2;10 |
| Emma | Danielle | F | 2;10 |
| Ellen | Joyce | F | 2;10 |
| Jennifer | Karen | F | 2;10 |
| Max | Alice | M | 2;11 |
| Charlie | Amy | M | 2;11 |
| Isabel | Mary | F | 3;0 |
| Heather | Tania | F | 3;1 |
| Kevin | Joanne | M | 3;1 |
| Kerry | Paula | F | 3;2 |
| Stephen | Donna | M | 3;2 |
| Jake | Liz | M | 3;2 |
| Becky | Amanda | F | 3;2 |
| Billy | Sarah | M | 3;2 |
| Lyle | Fran | M | 3;3 |
| Annie | Kimberley | F | 3;3 |
| Gareth | Carol | M | 3;3 |
| Lucy | Lesley | F | 3;4 |
| Ellie | Gail | F | 3;5 |
| Luke | Molly | M | 3;5 |
| Kieran | Ruth | M | 3;6 |
| Gus | Mandy | M | 3;9 |
| Oliver | Judy | M | 3;9 |
| Marie | Eileen | F | 3;10 |
| Izzie | Denise | F | 3;11 |
| Ella | Louise | F | 3;11 |
| Dan | Suzie | M | 4;2 |

contains 29 dyads aged between 2;10 and 4;2 (Table 1). Children by these ages are at a stage where full clauses are used, allowing for the analysis of morpho-syntactic variation such as agreement phenomena, negation and tense formation. To control the sample as much as possible, the participant selection process was guided by the following criteria: (i) both parents were born and raised in the community; (ii) the mother was the main caregiver, i.e., no substantial time was spent with extended family, childminders etc.; (iii) no child was in formal nursery education, to avoid effects of standardization.

The caregivers were provided with lightweight minidisc recorders (Sony MZ-R700) and lapel microphones (Sony ECM-T145) and requested to undertake a

series of recorded sessions with their child, in a variety of situations where interaction takes place. These included mealtimes, trips in the car, walks, and lots and lots of play. The recordings amounted to ten hours for each pair, 290 hours in total. The transcribed corpus totals ca.700,000 words (see detail in Smith et al. 2007).

3 Testing the acquisition of sociolinguistic norms

3.1 The variables

Extending our previous research, we selected six variables: two lexically restricted phonological variables (2a)–(2b); two lexical variables (2c)–(2d); two morphosyntactic variables (2e)–(2f). This allows us to test the acquisition of variation across a range of variables taken from different levels of the grammar.²

- (2) a. *We'll just go **r[ʌ]**nd by that **h[u:]sie**.*³
(Caregiver)
- b. *I **ca[n't]** see them. I **ca[na]** see them.*
(Ella 3;11)
- c. *I don't **know** it, I just dinna **ken**.*
(Caregiver)
- d. ***Aye** kissie kissie yes kissie **yes** huggle.*
(Oliver 3;9)
- e. *You're too good at **this** jigsaws. Are you finding all **these** edges?*
(Caregiver)
- f. *We'll go and see if the grapes **is** open, see if the grapes **are** ready.*
(Caregiver)

For each variable there is a binary distinction between a standard variant and a vernacular Scots variant. Across each of the variables we tested a range of measures which have been shown to have an influence on the acquisition of sociolinguistic norms. We detail these below.

² We do not include t/d deletion in this analysis (Smith et al. 2009). In this paper we limit the variables to those which have a binary distinction between a Scots variant and a standard variant.

³ Gloss: We'll just go round by that house.

3.2 Comparison with general community norms

CDS has been shown to be qualitatively different from adult to adult speech, with modifications at the lexical, syntactic and phonetic levels (e.g., Snow 1994). Foulkes et al. (1999, 2005) provide details of how it may also differ *quantitatively*: their study of word-medial intersonorant and word-final prevocalic (t) in Newcastle found much lower rates of the vernacular variant in CDS when compared to adult norms. Roberts (2002) finds similar results in a pilot study of the pronunciation of (ay) in words such as *kite* in Memphis, Tennessee.

We have at our disposal a 300,000 word corpus of adult to adult speech from the community (see detail in Smith 2000a). This provides a baseline from which we can compare the norms of CDS to the more general sociolinguistic norms in this variety. Our previous results showed that the *hooose* variable had much lower rates in CDS when compared to adult norms, while the Northern Subject Rule had very similar rates. In other words, quantitative differences in CDS are variable specific. To test this hypothesis further, we compared the overall rates in community norms with those in CDS across all six variables. The original corpus contains three generations of speakers, but in order to ensure comparability with the caregiver data, we use the youngest generation only (22–31 years old) in these analyses.

3.3 Caregiver input/child output

Generative models of first language acquisition are based on the premise that input has little bearing on output (e.g., Chomsky 1988). More usage-based accounts suggest otherwise (e.g., Tomasello 2003). Previous research on acquisition of variable forms shows a statistically significant correlation between caregiver and child rates of use for some variants (e.g., Kerswill and Williams 2005; Patterson 1992; Chevrot et al. 2011). Our 2007 results showed this to be variable specific, leading us to suggest that some variables may be acquired through a rule-based process while others rely on copying of input (see e.g., Westergaard 2009). We further test for the correlation between input and output both in overall distributions and across caregiver/child individual pairs in this larger database.

3.4 Age of child

Foulkes et al. (2005: 201) demonstrate that caregiver talk “gradually becomes more similar in character to that of the inter-adult mode as the children get older”

i.e., there are higher frequencies of the standard variant with caregivers in interaction with very young children but these are replaced by increasing frequencies of the nonstandard variant as the child gets older. Smith et al. (2007) showed that the variable which demonstrated social awareness showed age differentiation and the other did not. Again, we attempt to further these insights on age related use across this larger database.

3.5 Style shifting

The emergence of adult-like patterns of style shifting appear to be in place at least at pre-adolescence (e.g., Hoyle and Adger 1998; Cheshire 1978, 1982; Purcell 1984; Renn and Terry 2009) with evidence of meta-awareness of informal and formal variants also attested (e.g., Reid 1978; Romaine 1978, 1984; Chevrot et al. 2000). Studies which include a range of ages may provide further insights. Chevrot et al. (2000: 302) in their study of postconsonantal /r/ in French find “stylistic adaptation” in 10–12 year olds but not 6–7 year olds. Similarly, Kerswill and Williams’ (2000) study of the new town of Milton Keynes found more systematic style shifting with 12 year olds when compared to 8 and 4 year olds. Roberts (1994: 177) remarks that “the acquisition of social constraints on variation has its beginnings in early childhood, but the bulk of this learning appears to take place after the age of four”. It may be even younger still. Diaz-Campos (2001, 2005), for example, found Spanish speaking children in Venezuela aged 4;6–5;11 using patterns of style shifting which model the adult patterns, although the younger speakers aged 3;6–4;6 did not. Patterson (1992) also found style shifting in her 4 year old speakers, and suggests that this is the result of young children learning to associate standard variants automatically and unconsciously with certain types of interactions within the family environment. The “family environment” may be further circumscribed to the primary caregiver. Labov (2001: 437) states that style shifting in caregiver speech is one of the key criteria for children acquiring sociolinguistic norms in that “[l]inguistic variation is transmitted to children as stylistic differentiation on the formal/informal dimension [...] formal speech variants are associated by children with instruction and punishment, informal speech with intimacy and fun”. In Smith et al. (2007) we found style shifting with one variable, but not with another across this formal/informal dimension. To test for these constraints further, we divided the data into different stylistic contexts which naturally arise from the caregiver’s multifaceted roles in daily interaction with her child: teaching and play versus routine and discipline. Interactions where the context is unclear were excluded. In line with Labov (2001: 437), we hypothesize that the more formal contexts of teaching and discipline would have

higher rates of the standard variant and the more informal contexts of routine and play would have higher rates of the vernacular Scots variant.

3.6 Linguistic constraints

Chambers (2003: 174) suggests that there is no time gap between the acquisition of categorical rules and the acquisition of variable rules. However “complexity and conditioning” (e.g., Kerswill 1996: 199) of the variable in question might influence when linguistic constraints are acquired. For example, Kovac and Adamson (1981) find variable use of finite *be* deletion in their 5 year old African American Vernacular English speakers, but not the specific linguistic constraints generally associated with this feature (see also e.g., Guy and Boyd 1990). Further, Roberts (1994) demonstrates that some variable constraints within one variable may take longer to acquire than others: children as young as 3 years old in her sample had largely mastered the phonological constraints on t/d deletion, but only partly acquired the grammatical constraints. Our research on the same variable (Smith et al. 2009) mirrored these results. In contrast, Smith et al. 2007 showed a range of linguistic constraints on use – lexical and morphosyntactic – were acquired even by the very youngest speakers in the sample. We test whether this is the case across this wider range of variables by testing a number of linguistic constraints on use in the caregiver and child data.

3.7 The interaction of linguistic and social constraints, and the order of acquisition

Roberts (1994: 30) points out that most studies of acquisition tend to “focus on one type of constraint or the other”. By testing both linguistic and social constraints, we may be able to contribute to the debate on order of acquisition (e.g., Labov 1989). Labov’s (1989) research on (ing) in three children aged 4, 6, and 7 in Philadelphia leads him to suggest that social and stylistic constraints are acquired before articulatory and grammatical constraints. Patterson’s (1992) study of the same variable supports this: the 4 year olds in her study demonstrated adult-like patterns of style shifting but had not yet acquired the systematic patterns of use associated with grammatical class or discourse function. The older children in her sample (6 and 8 year olds) had acquired both. Roberts (1994), on the other hand, found that the children in South Philadelphia had mastered the linguistic constraints on t/d deletion, but had not yet acquired the social constraints on use. Youssef’s (1991: 96) longitudinal study of verbal marking in a

child (age 2;4 to 4;9) exposed to Trinidadian Creole and standard English suggest that social and linguistic constraints are acquired at the same time.

Our previous research suggests that order of acquisition is dependent on the particular variable being acquired (Smith et al. 2007). We explore which types of variables may lend themselves to acquisition of social constraints first, linguistic, or both at the same time.

In sum for each of the variables we test:

1. Comparison of overall distributions between community, caregiver and child.
2. Comparison across ages of children.
3. Correlations between caregiver/child pairs.
4. Comparisons across different situational contexts.
5. Comparisons across linguistic constraints.

We use each of these pieces of evidence as building blocks in uncovering patterns of acquisition.

4 Results

4.1 The *hoo* variable

To set the scene, we first revisit a variable which we have previously analyzed in detail in Smith et al. 2007: the alternation between the diphthong [ʌɪ] and the monophthong [u:] as in (2a), herein referred to as the *hoo* variable.

This lexically conditioned variable belongs to the MOUTH class (Wells 1982) of words, which have the orthographical form <ow> or <ou> as in *now*, *house* and *down*. [u:] is considered stereotypical of Scots or more northern varieties of English (e.g., Stuart-Smith 2003) and avoided in more formal speech (e.g., Eremeeva and Stuart-Smith 2003, Romaine 1984). It is a lexically diffused phonological variable, with the two variants being phonetically distinct.

4.1.1 Overall distributions

In a reduced sample in our adult to adult data, 94% are realized with the local variant. In fact, the standard variant is only used to a significant degree in the lexical item *how*. Table 2 shows the overall distributions for the caregivers and children. Note that with this and following the tables, the total number of contexts of use is indicated. Thus, for Table 2, there are a total of 6715 contexts of use for the *hoo* variable, and 37% of these are realized with the local form.

Table 2: Overall distribution of local form in the caregiver and child data (Total N = 9873)

| Caregivers | | Children | |
|------------|----|----------|----|
| N | % | N | % |
| 6715 | 37 | 3158 | 24 |

The caregivers use the local form 37% of the time, and the children use lower rates at 24%. Both of these percentages are considerably lower than that in the adult to adult data.

4.1.2 Linguistic constraints

Our 2007 paper demonstrated that certain lexical items were near-categorically standard: (*how* and discourse marker *now*) or near-categorically non-standard (*anow*, meaning *just now*), as in (3).

- (3) a. *H[ʌu] many sleeps till Christmas?*
(Lucy 3;4)
b. *N[ʌu], you're out to see grandma the day.*
(Caregiver)
c. *Mam fitt does that smell like an[u:]?*
(Becky 3;2)

Moreover, there was a statistically significant correlation between caregiver and child in the remaining variable lexical items. To test for this categorical vs. variable constraint in this larger dataset, Figure 1 shows the distribution of the local form by the most frequently occurring individual lexical items.

The figure shows that this result is replicated when the additional 18 speaker pairs are included – *how*, and discourse marker *now* are near-categorically standard and *anow* is near-categorically nonstandard in the caregiver speech and this is replicated in the children’s speech. Further, a Pearson’s Correlation reveals that there is a statistically significant matching between caregivers and children across the remaining variable lexical items ($r = 0.564$, $df = 13$, $p < 0.05$).⁴ This pro-

⁴ In the 2007 data, we noted the higher rates of the local form with the adverb *now* in the child data, but had “no explanation” at that time for the result. Note here that the caregivers and children have almost identical rates, thus we attribute the 2007 result to statistical fluctuation.

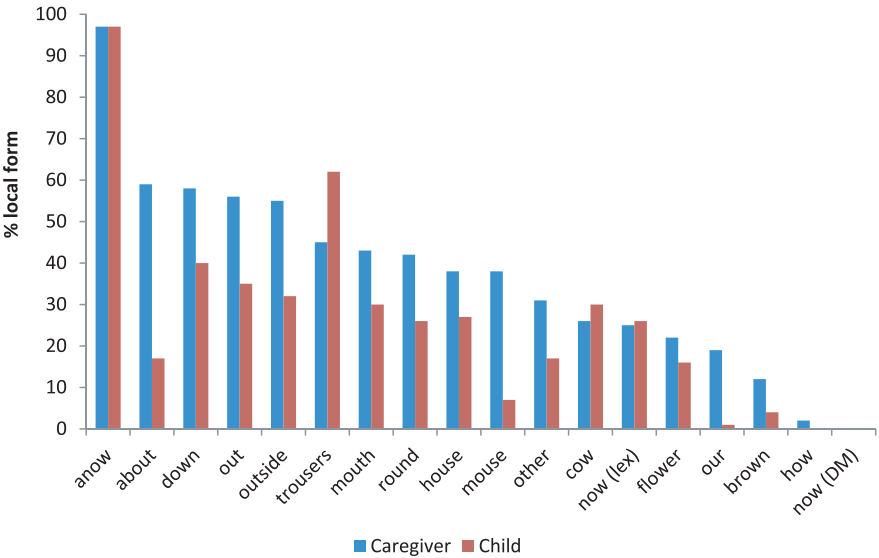


Fig. 1: Use of local form by frequently occurring lexical items

vides further support for our 2007 findings that lexically conditioned categorical and variable constraints in the input are matched in the output. We now remove these categorical contexts from further analysis.

4.1.3 Frequency by age

In our 2007 paper, we found that there was a difference in use of the local form by age: (i) younger speakers were near-categorical in their use of the standard form and (ii) the older the children were, the more they used the local form. Figure 2 divides the data by the 29 individual speaker pairs in order to test whether these patterns are maintained in these data. Note that in this and the following figures by individual speaker pairs, the children are ordered left to right from youngest to oldest.

Figure 2 shows that 10/11 of the younger speakers use lower rates of the local form when compared to the overall rates. The exception to this is one of the youngest speakers, Elizabeth, who has 33% use of the local form. Note too, that three of the older children – Gareth, Gus, and Izzie – also have low rates, despite belonging to the older cohort. For the caregivers, 7/11 caregivers of younger children have lower rates than the average compared to only 7/18 for caregivers of the older children.

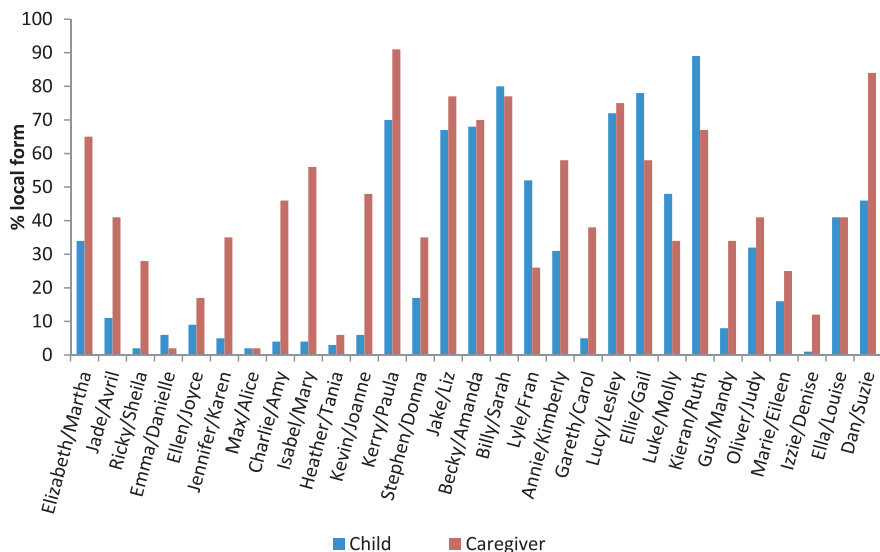


Fig. 2: Use of local form by individual caregiver/child pairs

We found that these results were similar across a number of other variables in this study, i.e., there was a split in percentages of use between children up to 3;2 (Kevin) and those above that age (Kerry onwards). In addition, a number of the older speakers had rates of use similar to the younger group. These descriptive statistics were further validated through a cluster analysis run using PASW 18. This revealed that the children fell into two groups, one comprising the eleven younger speakers and four of the older children (Gareth, Gus, Marie, and Izzie) and one comprising the other older children. Based on these findings, it might seem reasonable to group these four older speakers with the younger ones. However, although children differ in their rate of acquisition (e.g., Brown 1973), Pinker (1990) suggests that “[f]or just about every rule that has been looked at, 3-year olds obey it a majority of the time”. In our data we saw a marked difference in use of forms when comparing the 11 younger children to the 18 older ones: the younger cohort had a number of non-targetlike developmental forms but the older cohort had very few.⁵ Moreover, we show that the difference between these four older children when compared to their peers is probably due to caregiver input: the mothers of these children are more standard than the other caregivers

⁵ These include examples such as *dogs not bite* and *what’s piggies say?*

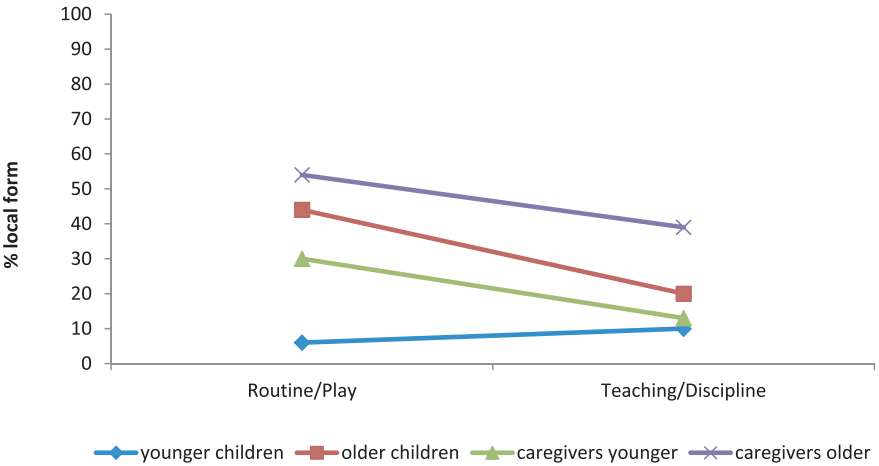


Fig. 3: Use of local form by situational context and age

of older children. For this reason, we group the children and caregivers into two broad categories: older vs. younger. Using these broad categories for this variable, t-tests reveal that the difference between the older and younger age groups is highly significant for the children ($p < 0.001$) (8% vs. 46%) and significant for the caregivers ($p < 0.05$) (31% vs. 52%). Thus, this larger dataset provides further support for the 2007 results regarding age of child: younger children (and their caregivers) use lower rates of the local form.

Finally, when the pairs are divided into the 11 younger children and the 18 older children, Pearson’s Correlation reveals statistically significant matching in terms of frequencies of use between the older children and caregivers ($r = 0.749$, $df = 16$, $p < 0.01$) but not with the younger children and caregivers ($r = 0.528$, $df = 9$, $p > 0.05$).

4.1.4 Stylistic constraints

Figure 3 shows how this variable patterns by situational context, divided into two broad categories of formal (Teaching and Discipline) and informal (Play and Routine). As the younger children have very different rates of use from the older children, we now separate the speakers accordingly.

Figure 3 shows style shifting across the different contexts for caregivers of both older and younger children. These patterns are replicated in the older

children. The difference between formal and informal contexts of use is statistically significant for these three groups (older children: chi square = 15.58, $df = 1$, $p < 0.001$, caregivers of younger children: chi square = 23.44, $df = 1$, $p < 0.001$, caregivers of old children: chi square = 16.08, $df = 1$, $p < 0.001$)⁶. Note however that for the younger children, there is no statistically significant difference for situational context (chi square = 1.1, $df = 1$, $p > 0.05$), and in fact the highest rates of use are in Teaching/Discipline, suggesting that these younger speakers have yet to acquire the patterns of systematic style shifting associated with this variable.

4.1.5 Summary of the *hoose* variable

For the *hoose* variable, our 2007 results are supported in this larger database. Younger children use the local form at much lower rates than the older children. Caregivers style shift with this variable, as do older children. Younger speakers do not. The categorical vs. variable lexical constraints on use in the input are replicated in the output, regardless of age.

4.2 Negation

The second variable involves different phonological realizations of negative particles in negative declaratives, both nonclitic and enclitic forms: local forms *na* and *nae* (realized as [na] and [ne] are in variation with the standard English realizations *n't* and *not* as in (2b).

There are a number of syntactic constraints on use across the different clause types (interrogatives, imperatives, tag questions and declaratives) but here we include declarative contexts only and concentrate on the phonological variants. The alternation is restricted to these negative particles thus it is not a productive phonological process in the dialect. The local variant is associated with working class speech and informal contexts (e.g., Macaulay 1991). It is also a stereotype in this dialect (Smith 2000a) as it is in Scots more generally. *Didna(e)* in particular, is singled out as a form that wouldn't be used in "polite speech" by the Edinburgh school girls recorded by Romaine (1984).

⁶ All chi-squares are carried out on nominal data. We chose to use chi-square tests rather than t-tests in cases where we examine situational context. As the children (and caregivers) had individually low numbers of tokens for the formal contexts of teaching and discipline, it would not have been possible to test for statistical significance otherwise.

Table 3: Overall distribution of local negation forms in the caregiver and child data (Total N = 8761)

| Caregivers | | Children | |
|------------|----|----------|----|
| N | % | N | % |
| 4845 | 72 | 3816 | 47 |

4.2.1 Overall distributions

In our adult to adult data, 99% of all negative contexts appeared with the local form. In other words, there was virtually no use of the standard form in vernacular data. Table 3 shows the overall distribution of forms in the caregiver and child data across these contexts of use.

The table shows that there are higher rates of the local form in the caregiver speech when compared to the *hooose* variable. However, the rates are still not as high as the adult to adult data. As with the *hooose* variable, the children have lower rates of the local form when compared to the caregivers.

4.2.2 Frequency by age

Figure 4 divides the data by individual speaker pairs.

Note that with this variable, three of the younger children use relatively high rates of the local form – Ellen, Jennifer and Max. Elizabeth, Charlie and Isabel are (near) categorically standard, as are Marie and Izzie at the opposite end of the graph. Despite this, t-tests reveal that the difference between the older and younger age groups (20% vs. 56%) is highly significant for the children ($p < 0.001$). It is not statistically significant for the caregivers however ($p > 0.05$) (63% vs. 80%). In other words, there is no difference in caregiver use according to whether their child is older or younger.

Furthermore, Pearson’s correlation tests reveal that there is a statistically significant matching in the older children/caregiver pairs ($r = 0.816$, $df = 16$, $p < 0.01$) but not in the younger pairs ($r = 0.102$, $df = 9$, $p > 0.05$). This is similar to the results to the *hooose* variable.

Figure 5 demonstrates that there is statistically significant style shifting with older children only (chi square: 20.99, $df = 1$, $p < 0.001$). In contrast to the rates of use, this looks very different to the *hooose* variable.

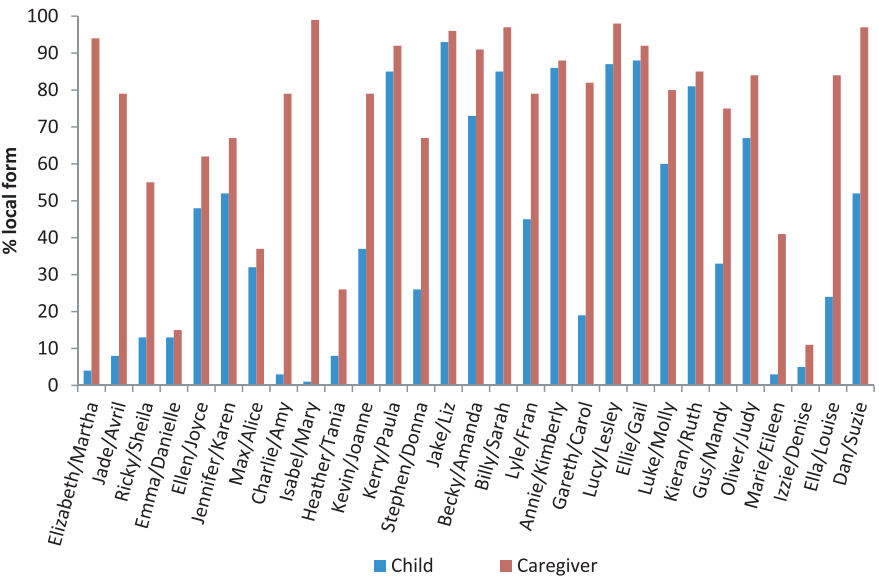


Fig. 4: Use of local negation forms by individual speaker pairs

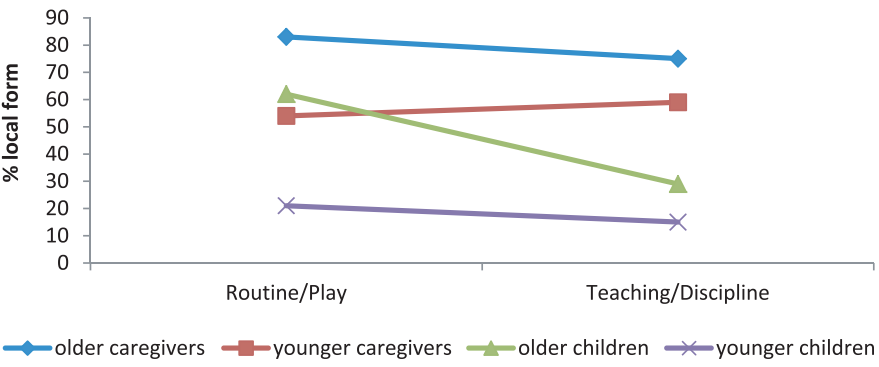


Fig. 5: Distribution of local negation forms by situational context

4.2.4 Linguistic constraints

Further investigation of the data showed that auxiliary *do* accounted for 31% of the total contexts of use. To test whether this very frequent context patterned differently from negation with other verbs (e.g., Tagliamonte 2006: 122), Figure 6

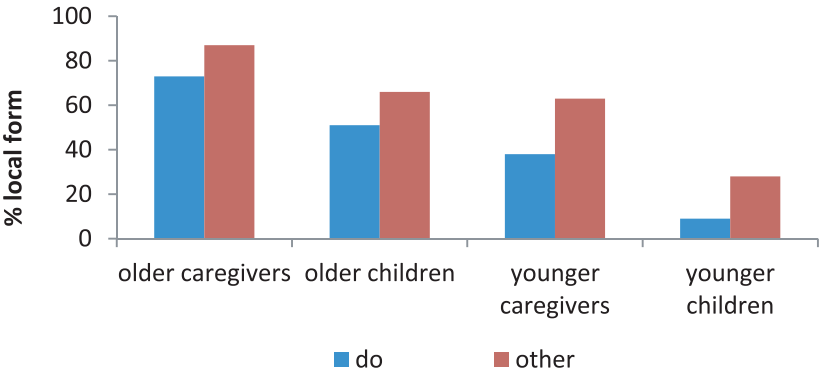


Fig. 6: Use of local form by negated verb and age

divides the data into *do* negation vs. negation with other verbs. The categorical speakers are removed from the data.

Figure 6 shows that caregivers and children in both age cohorts have the same pattern when the data are split into *do* and non-*do* negation: higher rates of the local form in non-*do* contexts. Thus, although the different age groups have different *rates* of use of the local form, on this linguistic measure they have the same *patterns* of use.

4.2.5 Summary of negation

As with the *hoose* variable, there is a statistically significant difference in rates of use between the younger and older children. All speakers show similar linguistic constraints on use, regardless of age. However, with stylistic constraints, only the older children style shift. We return to these points in the discussion.

4.3 *ken* vs. *know*

The first lexical variable we analyze is the alternation between *ken* and *know*, as in (2c). This form has been around since the 1300s (OED s.v. *ken*) but is still commonly used throughout Scotland (e.g., Miller 1993) and widely commented on as a feature of vernacular speech (e.g., Aitken 1979).

Table 4: Overall distribution of local lexical form *ken* in the caregiver and child data (Total N = 1898)

| Caregivers | | Children | |
|------------|----|----------|----|
| N | % | N | % |
| 1141 | 54 | 757 | 31 |

4.3.1 Overall distributions

In our adult to adult data, as with the local negation forms, *ken* is used near-categorically – 99%. Table 4 shows the overall distribution of use across the caregiver and child groups.

In the caregiver speech, there are considerably lower rates of use when compared to the adult to adult data and the children have even lower rates.

4.3.2 Frequency by age

When we looked at this variable across the individual speakers we found that 7 speakers – Ellen and Joyce, Stephen and Donna, Max, Ruth and Charlie – had fewer than ten tokens each. One child (Kieran) had no tokens at all. This is not surprising, as lexical variables are often the least frequently occurring in spoken data (e.g., Tagliamonte 2006: 64). Figure 7 shows the results for only those speakers who had >10 tokens⁷.

Note the split in the younger speakers vs. older speakers – with the younger speakers there is virtually no use of the local form at all. Not surprisingly, there is a statistically significant difference between the two age groups for the children ($p < 0.001$) (5% vs. 49%). On the other hand, despite a big difference in rates of use (37% vs. 63%), t-test reveals no statistically significant difference between the caregiver groups ($p > 0.05$).

For caregiver/child correlations, a Pearson’s Correlation reveals no statistically significant matching for the eight young child/caregiver pairs ($r = 0.554$, $df = 6$, $p > 0.05$) but for the older pairs there is ($r = 0.882$, $df = 14$, $p < 0.01$).

⁷ Note that some speakers have zero realizations of the nonstandard form e.g., Emma and Danielle. In these cases, the speakers had more than 10 tokens, but all of them were the standard variant.

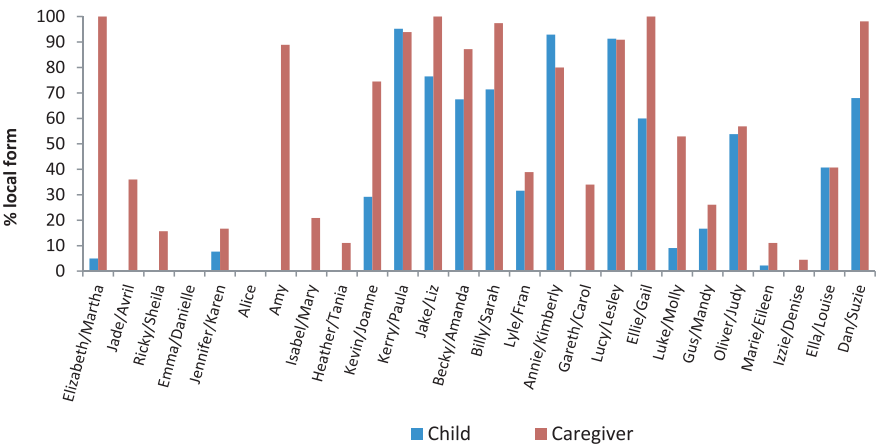


Fig. 7: Overall distribution of local form by speaker pairs

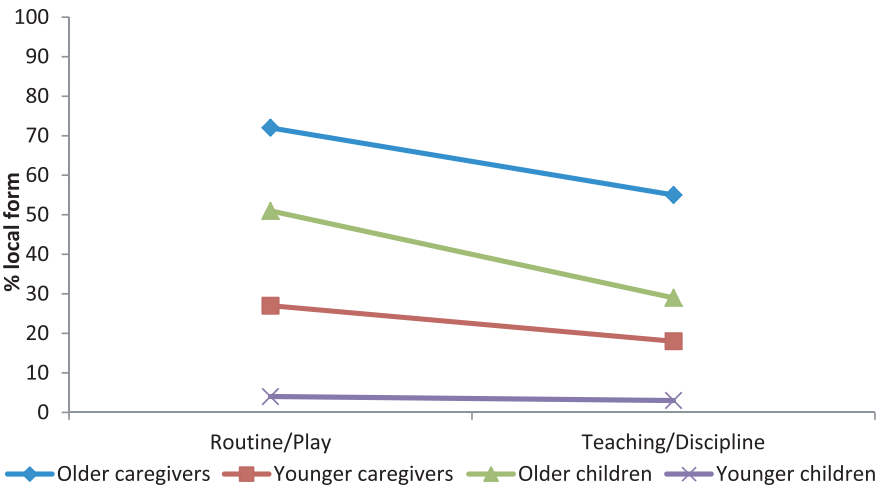


Fig. 8: Use of local form by situational context and age

4.3.3 Situational context

We now calculate the use of *ken* across situational context, dividing the speakers into younger and older cohorts.

Figure 8 shows that there is statistically significant style shifting for the older caregivers (chi square: 5.17, df = 1, $p < 0.05$) and children (chi square: 8.53, df = 1, $p < 0.01$). There is none for the younger children or caregivers.

4.3.4 Summary for *ken* vs. *know*

The caregivers use much lower rates of the local form compared to the adult to adult data. The younger speakers have very low rates of the local form and the difference in age groups is statistically significant. There is systematic style shifting with the older cohort but not with the younger in both caregiver and child speech.

4.4 *Aye* vs. *yes*

The second lexical variable is the alternation between the standard English form *yes* and the dialectal form *aye*, and in (2d). The standard form has been in use since Old English whereas the dialectal form entered the language around the beginning of the 17th century (OED sv. *aye*), and it remains today one of the most defining characteristics of the Scots tongue. In more formal situations, it is somewhat stigmatized: for example, it is reported that pupils are reprimanded for using it in school (e.g., Matheson and Matheson 2000: 217). It is one of the forms cited by schoolchildren in Romaine (1984), which wouldn’t be used in “polite” speech. In other words, it is a stereotype.

4.4.1 Overall distributions

There is near categorical use of the local form in adult data: 99% use of *aye*. Table 5 shows the overall distribution of use for the local form across caregiver and child speech.

Although lower than the adult to adult data, the caregivers still have high rates of the local form – 80%. The children are considerably lower, at 53%.

Figure 9 shows the use of *aye* by individual speaker pairs.

As with *ken*, there is a split in use of the local form, with generally much lower rates amongst the younger children. 5/11 children have in fact no use of *aye*

Table 5: Overall distribution of local form *aye* in the caregiver and child data (Total N = 8106)

| Caregivers | | Children | |
|------------|----|----------|----|
| N | % | N | % |
| 3202 | 80 | 4904 | 53 |

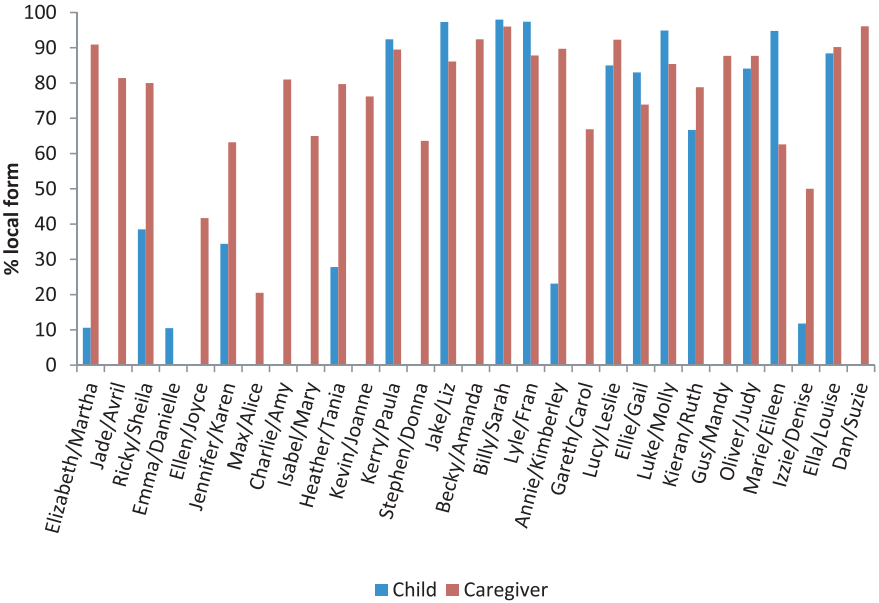


Fig. 9: Overall distribution of local form by speaker pairs

at all. Note too, however, that 4 of the older children are also near categorical. Despite this, there is a statistically significant difference in the age groups across the children ($p < 0.01$) (11% vs. 56%) and the caregivers ($p < 0.05$) (61% vs. 82%).

Pearson’s correlations of the young and old child/caregiver pairs show no statistically significant matching of frequencies (young pairs: $r = 0.206$, $df = 9$, $p > 0.05$, older pairs: $r = 0.254$, $df = 16$, $p > 0.05$). This is most probably due to the fact that a number of the older children categorically use the standard where their mothers are variable.

4.4.2 Stylistic constraints

Figure 10 shows the use of *aye* across three different situational contexts, with the younger and older pairs separated. The (near) categorical speakers have also been removed.

For the variable speakers, the figure shows that there is no style shifting for the caregivers and children in the older cohort. The lines are flat, with no statistically significant difference across contexts the contexts for children (chi square:

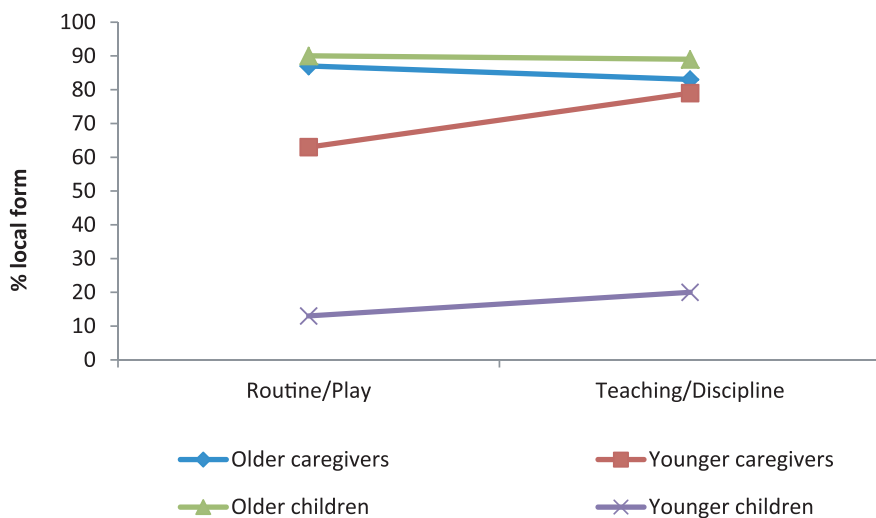


Fig. 10: Use of local form by situational context and age

0.15, $df = 1$, $p > 0.05$), or caregivers (chi square: 1.12, $df = 1$, $p > 0.05$). In contrast, for the younger cohort, the patterns are very different; the highest rates of the local form appear in the formal contexts for both caregivers and children, in other words, the opposite of the expected pattern of higher rates of the local form in informal contexts. The difference is statistically significant for the caregivers (chi square: 5.76, $df = 1$, $p < 0.05$), but cannot be tested for the children due to too few contexts of use.

4.4.3 Summary

There is a highly significant difference between older and younger children and their caregivers in the use of the local form. There is no style shifting with the older children and caregivers, and the opposite pattern to what is expected for younger children and caregivers. We return to these points later.

4.5 Plural demonstratives

The first morphosyntactic variable we consider is the use of “singular” demonstratives in plural contexts. In this dialect and dialects in the north east of

Table 8: Overall distribution of local forms in the caregiver and child data (Total N = 1276)

| Caregivers | | Children | |
|------------|----|----------|----|
| N | % | N | % |
| 810 | 80 | 466 | 68 |

Scotland more generally, singular demonstratives proximate *this* and distal *that* can be used with plural subjects (e.g., Macafee and O’ Baoill 1997; McRae 2000),⁸ as in (2e). In contrast to the *hoose*, negation and the lexical variables, the local forms are not overtly commented on by the speakers of these dialects, despite their widespread use (e.g., McRae 2000; Smith 2000).

4.5.1 Overall distributions

In our adult data, the local form is used 84% of the time. Table 8 shows the overall distribution of the local form in the caregiver and child data.

The table shows that in sharp contrast to the previous 4 variables, the caregivers use the local form at the same rates as the adult to adult data. The children’s rates are slightly lower.

4.5.2 Frequency by age

Figure 11 shows the results by individual speaker pairs.

The figure shows that all caregivers and all children have high rates of the local form, regardless of age. There is no statistically significant difference between the older and younger children ($p > 0.5$) (74% vs. 75%) or the caregivers ($p > 0.1$) (76% vs. 85%). A Pearson’s correlation for caregiver/child pairs reveals a statistically significant correlation ($r = .379$, $df = 27$, $p < 0.05$).⁹

⁸ A third variant *thon* is used in distal demonstratives but it does not appear in the caregiver/child data, hence it is not considered here.

⁹ A reviewer has suggested that nonagreement in this case may be “easier” at the cognitive level. The widespread use of this form in the adult data precludes such an interpretation.

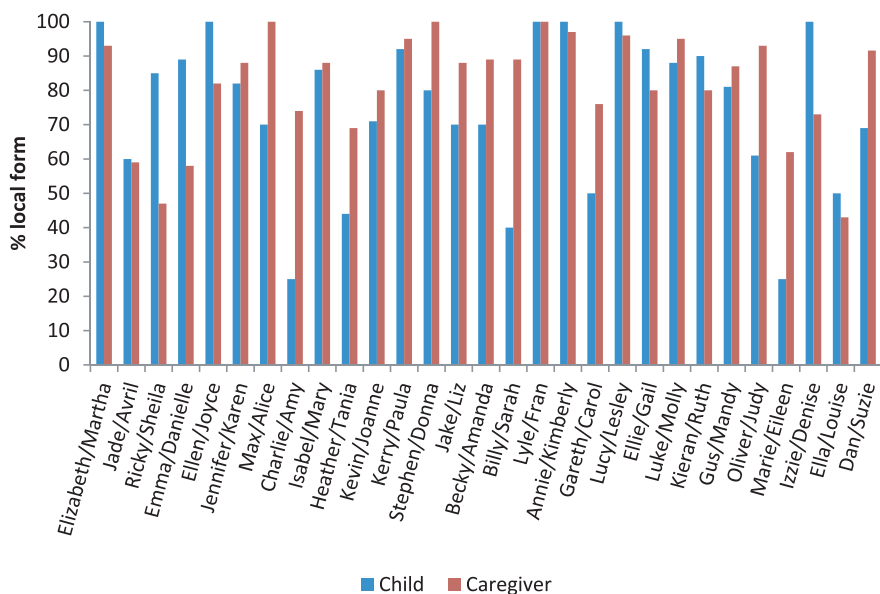


Fig. 11: Use of local form across individual speaker pairs

4.5.3 Social constraints

Figure 12 looks at the use of this variable across situational context in caregiver and child speech. Due to the similarity of results in Figure 14, we do not separate the speakers by age.

As Figure 12 shows, there is no systematic style shifting in the use of *this* and *that* with plural demonstratives according to situational context. In fact, for the caregivers and children, the lowest use of the local form appears in informal contexts and the highest in formal ones, although this is not statistically significant ($p > 0.05$ in both cases).

4.5.4 Linguistic constraints

In the adult to adult data, we find a number of internal constraints on use. Specifically, there are higher rates of the local form in distal (*that/those*) contexts, as opposed to proximal (*this/these*) contexts. We test these constraints in Figure 13.

The figure shows a more to less hierarchy of use across the proximal and distal contexts in caregiver speech and this is replicated in the child speech. Across both sets of speakers, this is highly statistically significant ($p < 0.001$).

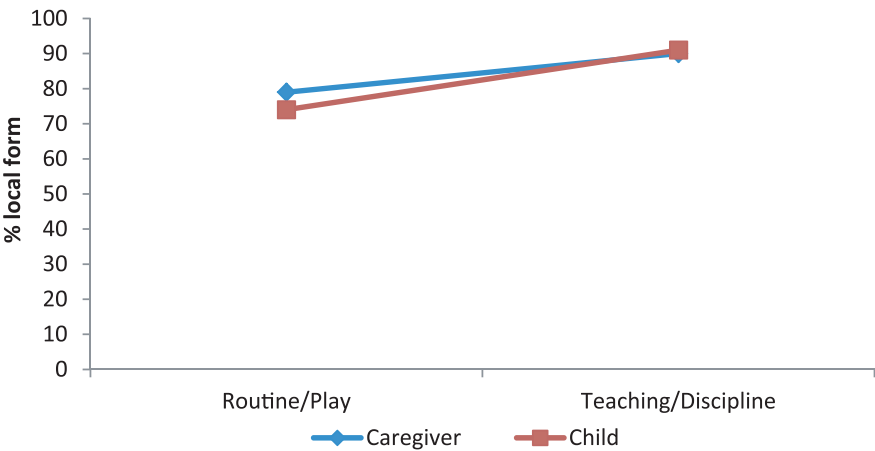


Fig. 12: Use of local form by situational context in caregiver and child speech

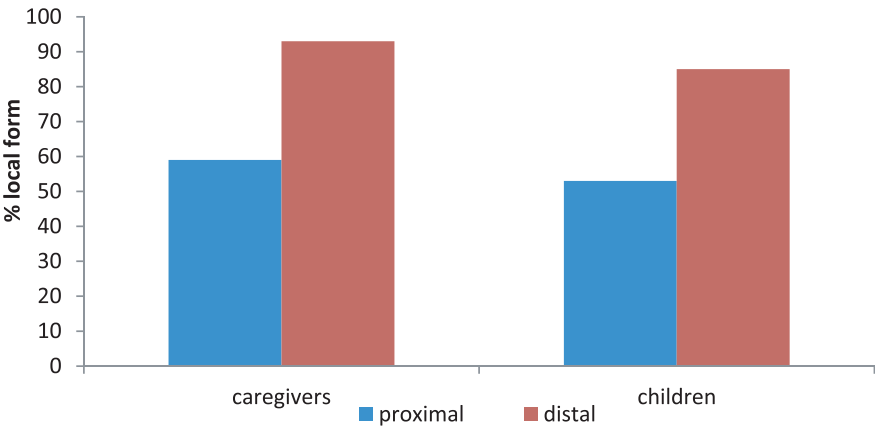


Fig. 13: Use of local forms by proximal and distal contexts

4.5.5 Summary of plural demonstratives

In sum, for plural demonstratives, caregivers have rates of use similar to adult to adult use. All children have high rates of the local form, regardless of age. There is no style shifting, but there are linguistic constraints in caregiver speech which are modeled in the children’s speech.

4.6 Northern Subject Rule

Finally, we revisit another variable discussed extensively in our 2007 paper – the use of -s in 3rd person plural contexts, otherwise known as the Northern Subject Rule (NSR), as in (2f). In standard English, -s in the present tense appears on 3rd person singular contexts only. In Buckie, as well as in other varieties of Scots, the -s inflection can appear in 3rd person plural full NP contexts but not with pronominal *they* (e.g., Tagliamonte et al. 2005). Despite its widespread use, dating back to the 13th century (e.g., Murray 1873) to the best of our knowledge it is not overtly commented on by speakers of Scots. It is certainly not the subject of comment on in the community in question. In fact, speakers are surprised when this form is pointed out to them (Smith 2000a), suggesting that it has little or no “social recognition”.

4.6.1 Overall distributions

In the adult to adult data, we found a categorical vs. variable split in use of -s as the historical record predicts: in *pro they* contexts, less than 1% appear with -s while full NPs appear 60% of the time with -s.¹⁰ In our 2007 paper, we found that this categorical vs. variable split was evident not only in the caregiver data, but in the children’s too. Table 9 shows the results across the current corpus.

The table shows that the split is maintained across this larger dataset. In other words, from the very beginning, the children have this categorical vs. variable rule. We now remove the categorical contexts of pronominal *they* from the remaining analysis.

Table 9: Overall distribution of -s in 3rd person plural contexts (Total N = 3815)

| Caregivers | | | | Children | | | |
|------------|---|------|----|----------|---|-----|----|
| pro they | | NP | | pro they | | NP | |
| N | % | N | % | N | % | N | % |
| 1365 | 1 | 1342 | 72 | 480 | 1 | 628 | 70 |

¹⁰ The tokens that did have *they* + -s (N = 7) were specialized to *verbs of communication* which Leech (1987: 11) points out have their own specialized system connected to the semantics of these verbs.

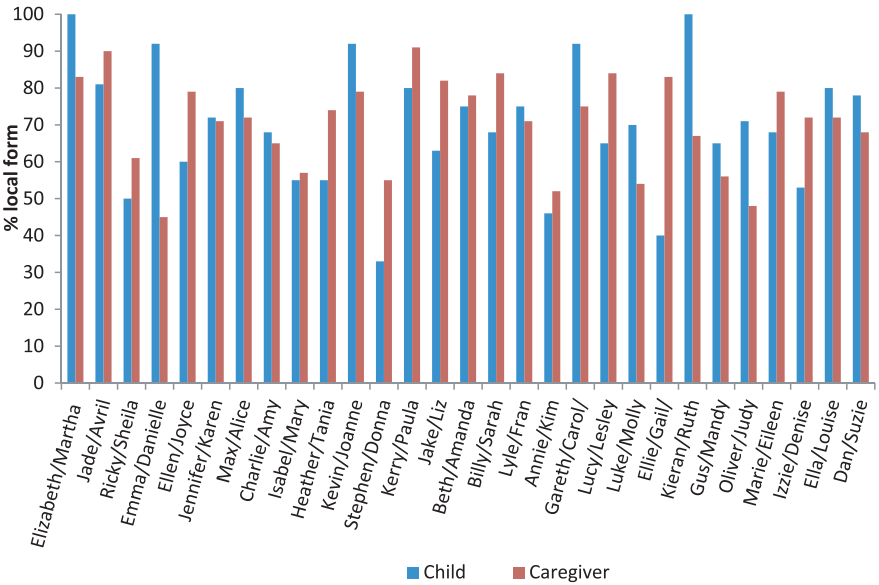


Fig. 14: Use of the local form across individual speaker pairs

4.6.2 Frequency by age

Figure 14 shows the data across the 29 speaker pairs.

The figure reveals that all caregivers and all children – regardless of age – have high rates of the local form. There is no statistically significant difference between young and old for the children ($p > 0.05$) (73% vs. 68%) or their caregivers ($p > 0.05$) (71% vs. 71%). In addition, there is no correlation between child and caregiver pairs ($r = 0.202$, $df = 27$, $p > 0.5$).

4.6.3 Linguistic constraints

In addition to the NP vs. pro split, a number of other contexts were shown to be (near) categorical in the adult data (Smith 2000) and these were replicated in the caregiver/child data in the 2007 paper. These included existentials as in (4a) “singular” pronominal demonstratives *this* and *that*, as in (4b) and the demonstrative pro *here* (4c). Pro *these* and *those* (4d)–(4e) were near-categorically standard. Table 10 shows the distribution of these contexts across the 29 speaker pairs.

Table 10: Distribution of -s across a number of 3rd person contexts

| | Caregivers | | Children | |
|--------------------|------------|-----|----------|----|
| | N | % | N | % |
| existentials | 454 | 94 | 197 | 96 |
| <i>this/that</i> | 200 | 100 | 74 | 99 |
| <i>here</i> | 34 | 100 | 14 | 93 |
| <i>these/those</i> | 13 | 0 | 22 | 9 |

- (4) a. *Ok **is there** any deers?*
(Gus 3;9)
- b. ***This is** hard ains.*
(Jake 3;2)
- c. ***Here's** more berries in the hedge.*
(Caregiver)
- d. *Wow **these are** good are they?*
(Marie 3;10)
- e. ***Are those** mine?*
(Becky 3;2)

Table 10 shows that the near-categorical contexts in our 2007 paper are also categorical in this larger database.¹¹ We now remove these (near)-categorical contexts for the following analyses.

4.6.4 Situational context

Our 2007 paper showed that there was no style shifting with this variable. Figure 15 shows the results for this larger database.

The figure shows that there is style shifting in the caregiver and child speech, but in the opposite direction to what is expected: Teaching/Discipline has the highest rates of use for both groups, although the effect is not statistically significant ($p > 0.05$ in both cases).

¹¹ We note that in these contexts, grammaticality judgments show that both variants are acceptable. With *pro they*, however, -s is rejected outright by all speakers, with the exception of *verbs of communication* (see Adger and Smith 2010). The first author, a native speaker of this dialect, has the same judgments.

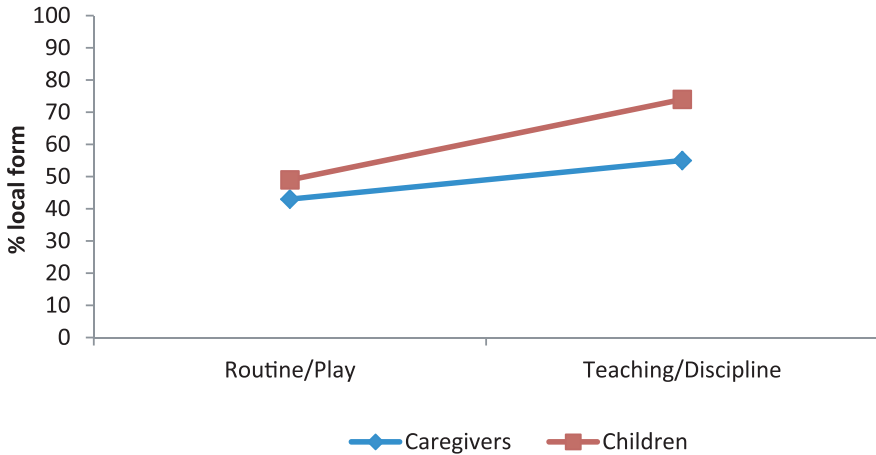


Fig. 15: Use of local form across situational context

4.6.5 Further analysis of linguistic constraints

In addition, in fully variable contexts, we found a wide range of percentages of local use across a number of other configurations in the 2007 paper: interrogatives (5a) vs. declaratives (5b), *be* (5c) vs. other lexical verbs (5d).

- (5) a. *And farr **is** mine?*
(Becky 3;2)
- b. *Pirates **are** bad mannies.*
(Caregiver)
- c. *My legs **is** not very long.*
(Heather 3;0)
- d. *No muppets go across the bridge.*
(Ellen 2;10).

We sum up this variable by graphing the use of the local form across the different linguistic contexts of use discussed above. Figure 16 shows the results.

Figure 16 shows remarkable similarities across the different contexts of use in caregiver and child speech. We return to this point below.

4.6.6 Summary of Northern Subject Rule

In sum, there is no style shifting in the expected direction with *-s* in 3rd person plural contexts, but the categorical vs. variable linguistic constraints in the

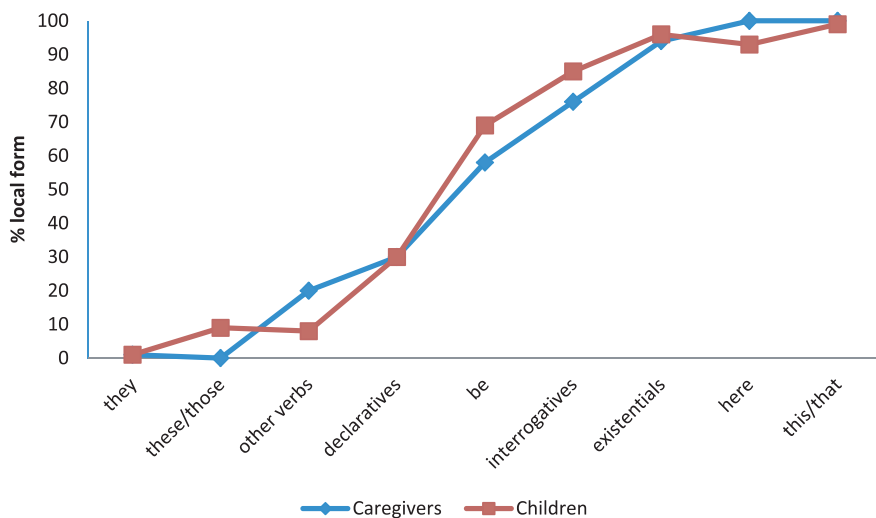


Fig. 16: Use of the local form across a number of 3rd person plural contexts

caregiver data are replicated in the child data, as are the fully variable contexts of use.

5 Discussion

We have investigated six linguistic variables in the speech of caregivers and children in Buckie. What do these results tell us about the acquisition of variation? At the beginning of the paper we laid out a number of measures by which we would assess the mechanisms involved in this process. Specifically, comparison in both input (caregiver speech) and output (child speech) with 1) adult to adult norms 2) different age groups 3) different situational contexts 4) different linguistic contexts. We review each of these in turn.

5.1 Comparison with adult to adult norms

Across the six variables we reviewed, we found different rates of use of the vernacular forms between community, caregiver and child. These are summarized in Figure 17.

Smith et al. (2007) demonstrated that frequencies of use of vernacular variants in caregiver speech differed from community norms with the morphosyntac-

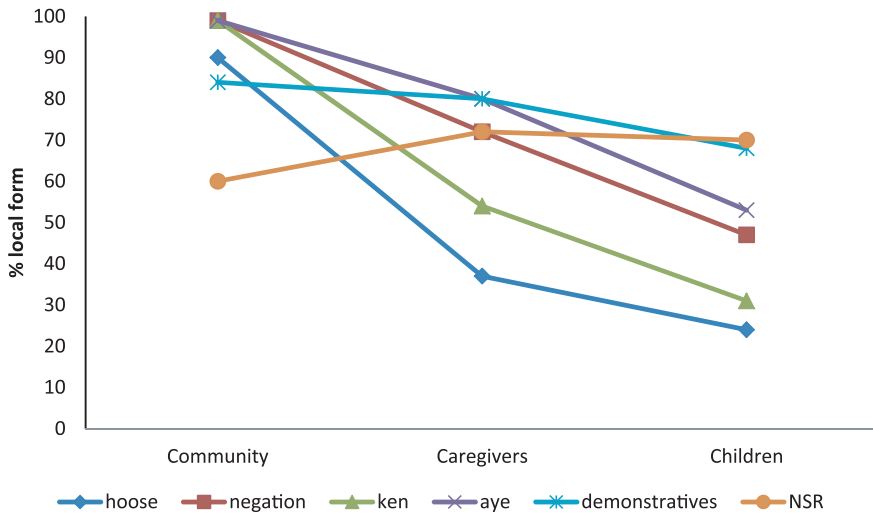


Fig. 17: Use of local form in community, caregivers and children

tic variable but not with the lexically conditioned phonological variable. These initial overall frequencies of use lend further support to this finding. Figure 17 demonstrates that for the *hooose* variable, negation, *ken* and *aye*, the caregivers used considerably lower rates of the local form when compared to adult to adult patterns of use. We also note that the children have lower rates than the caregivers across these four variables. With plural demonstratives and the NSR, on the other hand, the caregivers have very similar rates, or in the case of the NSR, even higher rates than general community norms. Moreover, the children replicate these overall distributions. Thus, just as in our 2007 paper, there is an initial split between the morphosyntactic variables, and the remaining four lexical/phonological variables.

However, this larger study reveals finer distinctions within these broad categories now that we have six variables at our disposal. Within the lexical category, the caregivers use relatively high rates of *aye* (80%) when compared to *ken* (54%). The same is true of the children. The lexically conditioned phonological variables are the same: the caregivers and children have higher rates of the local form with negation when compared to the *hooose* variable. This, despite the fact that these four variables are near-categorical in the adult to adult data. Part of this complexity may arise from the different age cohorts, where variants may be treated differently in CDS depending on the child's age (Foulkes et al. 2005). It was to this question that we next turned.

5.2 Frequencies of use according to age

When we divided the children and caregivers into individual speaker pairs we found that with some variables there was a statistically significant difference between (caregivers of) younger and older speakers: (caregivers of) older speakers had higher rates of the local form. With other variables there was no difference. Table 11 summarizes these results.

The table shows that with the *hoose* variable, negation, *ken* and *aye*, there are statistically significant differences in rates of the local form with the younger children when compared to the older children, with lower rates of use in the younger cohort. This is found for *hoose* and *aye* in the caregivers as well. For demonstratives and the NSR, there is no difference in either caregivers or children. In other words, some variables are used at higher/lower rates according to the age of the child and others are not.

These findings across the different age groups may be linked to input. Table 12 summarizes whether there is a statistically significant correlation between caregivers and children in rates of use of the local forms.

Some general tendencies emerge in Table 12. There is no correlation in frequencies of use for the NSR across all children. This is what we found in our 2007

Table 11: Statistically significant differences in frequencies of use between older and younger age groups

| | Children | Caregivers |
|----------------|----------|------------|
| <i>hoose</i> | ✓ | ✓ |
| negation | ✓ | X |
| <i>ken</i> | ✓ | X |
| <i>aye</i> | ✓ | ✓ |
| demonstratives | X | X |
| NSR | X | X |

Table 12: Statistically significant correlations between caregiver input and child output

| | Younger children | Older children |
|----------------|------------------|----------------|
| <i>hoose</i> | X | ✓ |
| negation | X | ✓ |
| <i>ken</i> | X | ✓ |
| <i>aye</i> | X | X |
| demonstratives | ✓ | ✓ |
| NSR | X | X |

paper also. For demonstratives, there is a correlation. For the remaining variables, we find that there is a statistically significant correlation with the older cohort of speakers, but not with the younger. The exception to this is *aye*, where there is no correlation with either age group.

How can these results be explained? Vosoughi (2010: 3) finds that as a child gets closer to the age of acquisition of particular words, the “caregivers tune the prosodic and distributional characteristics of their speech to the linguistic ability of the child”, using, for example, exaggerated prosody with certain lexical items. This process is then “reversed” as the child passes the age of acquisition, with e.g., “normal” prosody being employed from there on in. This is said to demonstrate “a new kind of adaptive behavior by the caregivers in the context of child language development” (Vosoughi 2010: 4). The results for the current research may also show a new kind of adaptive behavior, but with an added local/standard twist: the caregivers exaggerate the use of the standard variant with younger children but once the children pass the age of acquisition, then the caregivers revert to adult-like norms in their speech (see also Smith et al. 2007).¹² Thus, although the caregivers of the younger children are in fact using both variants, the standard one may have more prominence in the speech of the caregiver and this is the form that the child will hear, and use, first. This would explain the lack of correlation in this younger age group: in some cases their use is categorically standard, despite robust variation in the caregiver speech.¹³ This is demonstrated in the interaction between Amy and Charlie aged 2;11 (6). Amy’s last negative utterance – which is the standard form – is louder and longer than the preceding local forms:

- (6) (Charlie) *Abbey didn’t show us that thing to me.* (Amy) *Did she nae?* (Charlie) *I want to go and visit Abbey.* (Amy) *Well we’ll be gan to Brewsters in a little while darling. See Abbey later.* (Charlie) *I want to go and see Abbey this minute.* (Amy) *But Charlie you **havana** had your shower. And you **havana** got your clothes on.* (Charlie) *I want to go and see her. I want to see her mam.* (Amy) *Charlie I **dinna** think she’s up. It’s really early.* (Charlie) *Well we’ll go-.* (Amy) *Abbey **winna** be up.* (Charlie) *Can I go and see Kelly?* (Amy) *Darling they **winna** be up yet. They won’t be up.*

¹² We also note that 4 of the older speakers – Gareth, Gus, Marie and Izzie – have consistently lower rates of the local forms when compared to their peers. However, their caregivers also have lower rates.

¹³ We note that there is no correlation between older child and caregiver with *aye*. This may be due to the fact that five of the older children are categorically standard with this variable.

This privileging of the standard form through emphatic use might explain why Charlie uses the local form only 3% of time, despite the fact that his mother, Amy, uses it nearly 80% of the time.¹⁴ Thus we have an added dimension in acquisition: certain standard variants are exaggerated in caregiver input and this profoundly affects child output.

In terms of the overall findings for input/output, the (lack of) correlations across the different variables lead us to believe that some variables may be learned on a word by word basis i.e., through a copying of surface forms (e.g., Chevrot et al. 2000) e.g., *hooose*, *ken*, whereas others such as NSR involve a more complex array of constraint acquisition (e.g., Labov 1989; Roberts 1997a; Smith et al. 2009; Henry et al. 1997a; Westergaard 2009). Some may involve both (e.g., Patterson 1992).

However, frequencies of use are not the only clue to the processes at work in the acquisition of variation. Patterns of use – both linguistic and sociolinguistic – may provide further clues to how and when variable patterns are acquired.

5.3 Linguistic constraints

Chambers (2003: 173) suggests that variable rules are acquired at the same time as categorical rules. Our results support this hypothesis. Table 13 reviews the linguistic constraints on use across four of the variables where these were tested in our data.

The table shows that all children, regardless of age, replicate the linguistic patterns of use in the caregiver data. In other words, they have acquired the complex linguistic constraints on use, both categorical and variable, at least in those we tested for. For example, with the *hooose* variable, there were categorical constraints on lexical items *how*, discourse marker *now*, and *anow* and a statistically significant correlation between caregivers and children for the remaining variable lexical items.¹⁵ For *negation*, there was a statistically significant hierarchy of use according to verb type in both caregiver and child speech: lower rates of local form with *do* negation when compared to other verbs.

As a child has more or less mastered his/her first language by the age of three (e.g., Brown 1973), these results, for these variables at least, suggest that variable

¹⁴ It would also explain the somewhat anomalous result for the *hooose* variable, where children generally acquire monophthongs before diphthongs but here the younger children acquire the standard diphthong variant first.

¹⁵ Note that the younger children have a tick *and* a cross for this variable: this is because they had acquired the categorical rule for *how*, *now* and *anow* but unlike general community or caregiver norms, were near-categorical in their use of the standard form in the remaining variable contexts.

Table 13: Linguistic constraints by age and variable

| | Younger Caregivers | Children | Older Caregivers | Children |
|----------------|--------------------|----------|------------------|----------|
| <i>hoo</i> | ✓ | ✓/X | ✓ | ✓ |
| negation | ✓ | ✓ | ✓ | ✓ |
| demonstratives | ✓ | ✓ | ✓ | ✓ |
| NSR | ✓ | ✓ | ✓ | ✓ |

Table 14: Style shifting by age and individual variable

| | Younger Caregivers | Children | Older Caregivers | Children |
|----------------|--------------------|----------|------------------|----------|
| <i>hoo</i> | ✓ | X | ✓ | ✓ |
| neg | X | X | X | ✓ |
| <i>ken</i> | X | X | ✓ | ✓ |
| <i>aye</i> | X | X | X | X |
| demonstratives | X | X | X | X |
| NSR | X | X | X | X |

constraints are acquired at the same time as categorical ones in these pre-school years. Thus, although *rates* of use may vary, the *patterns* are the same. Our results also question Johnson’s (2005) assertion that variable input causes a delay in the comprehension of grammatical morphology (see also Yang 2002; Miller 2007; Miller and Schmidt 2009). Although we do not test here for comprehension of forms per se, there is no delay in the systematic *production* of these forms in line with adult norms, as evidenced in the mastery of constraints for the NSR and demonstratives.

However, not all linguistic constraints may be acquired this early: our research on t/d (Smith et al. 2009) showed acquisition of articulatory, but not grammatical constraints in the same database of speakers, highlighting Kerswill’s (1996: 199) point regarding the linguistic complexities of the variable in question. We are currently exploring this further through the analysis of a number of other variables.

5.4 Stylistic constraints in interaction with age

Table 14 provides a summary of the variables which show systematic style shifting (✓) and those which show either no statistically significant difference in use across the informal/formal dimension or pattern in unexpected ways, i.e., higher rates of the vernacular form in more formal contexts (X).

Table 14 shows that there is no style shifting with the younger children across any of the variables. However, we note that with the exception of *hooose*, there is no style shifting in the caregivers of these children either. With the older cohort, there is style shifting with *hooose* and *ken* in caregivers and children, and with the children also for *negation*. Thus, a general picture emerges that when there is style shifting in the caregiver speech, there is also style shifting in the child's speech. When there is no style shifting in the caregiver speech, there is none in the child's either. This provides further support for our 2007 findings which suggested that style shifting is input-based. Moreover this may be maturational. Chevrot et al. (2011) suggest that for adult-like patterns of variation to appear, the child must be exposed to the cumulative effect of the input. This may explain why the youngest children are less likely to pattern like their mothers, whereas the older children are more likely to follow parental norms.

Why is there a general lack of style shifting overall with these variables? We propose that there are different explanations for the different age cohorts. For the caregivers of the younger children, their lack of style shifting may arise from over-monitoring in CDS with particular variables, which result in unsystematic patterns of use. Once there is less monitored speech, as Foulkes et al.'s (2005) results suggest, a more systematic pattern emerges, as is shown in the older age group's use. For the older cohort, there is style shifting with *hooose* and *ken* only. These are exactly the variables which showed much lower rates of the local form compared to community norms (Figure 16). Thus, when there is a window of variation, this older cohort put the variants to good use in CDS. In contrast, with *aye* and *negation*, CDS overall rates were nearer to community norms, and their patterns of use are too. As in the adult to adult data, there is no style shifting.

The results across situational context and age suggest that some variables come to mirror community norms very quickly while others remain quite different in these early stages. This leads us to believe that these forms may have different “sociolinguistic value” in CDS at the start of the acquisition process, but these come to mirror community values more generally, where the local variants are the norm. This leads us to predict that the local variants will come to dominate caregiver/child interaction in the home environment as the child gets older: the caregiver's, and hence, children's rates of use will reach the same level as the general community norms i.e., used categorically. The standard variants, now learned, will be consigned to more formal situations such as school and talking to outsiders.¹⁶

16 We also predict that the morphosyntactic variables will remain at the same rates of use, regardless of the context. These are not amenable to higher or lower use across different contextual styles.

5.5 The order of acquisition

Finally, recall the debate on the order of acquisition, where some studies show that social constraints are acquired first (e.g., Labov 1989) while others show that these are acquired after linguistic constraints (e.g., Roberts 1994). Where we tested for both social and linguistic constraints, our results showed that for the majority of the variables, there was no style shifting, in either caregiver or child speech, thus it is not possible to use these as evidence for order of acquisition. The only test case is with *hooose*: the younger children had acquired the categorical lexical constraints on use, but did not show systematic patterns of style shifting evidenced in their caregivers' speech. However this is only one piece of evidence. We now need to explore a series of variables which demonstrate social and linguistic constraints to uncover what is acquired when, thus we leave this question for future research.

6 Conclusion

We now return to our original question posed in the introduction: when do children acquire the highly complex patterns of variation, both linguistic and sociolinguistic, widely attested in adult speech? We conclude that they do so at a very early age, in tandem with language acquisition more generally. However, the details of *how* the acquisition process unfolds may differ according to the variable under study and even then this may be “different for different social groups” (Kerswill and Williams 2002: 101).

References

- Adger, David & Jennifer Smith. 2010. Variation and agreement: A lexical feature-based approach. *Lingua* 120(5). 1109–1134.
- Aitken, A. J. 1979. Scottish speech: a historical view, with special reference to the Standard English of Scotland. In A. J. Aitken and Tom McArthur (eds.), *Languages of Scotland*, 85–118. Edinburgh: Chambers.
- Brown, Roger. 1973. *A first language*. Cambridge, MA: Harvard University Press.
- Chambers, Jack. 2003. *Sociolinguistic theory*. Oxford: Blackwell.
- Cheshire, Jenny. 1978. Present tense verbs in Reading English. In Peter Trudgill (ed.), *Sociolinguistic patterns in British English*, 52–68. London: Edward Arnold.
- Cheshire, Jenny. 1982. *Variation in an English dialect: A sociolinguistic study*. Cambridge: Cambridge University Press.

- Chevrot, Jean-Pierre, Laurence Beaud & Renata Varga. 2000. Developmental data on a French sociolinguistic variable: Post-consonantal word-final /R/. *Language Variation and Change* 12. 295–319.
- Chevrot, Jean-Pierre, Aurélie Nardy & Stephanie Barbu. 2011. Developmental dynamics of SES-related differences in children's production of obligatory and variable phonological alternations. *Language Sciences* 33. 180–191.
- Chomsky, Noam. 1988. *Language and problems of knowledge*. Cambridge, MA: MIT Press.
- Díaz-Campos, Manuel. 2005. The emergence of adult-like command of sociolinguistic variables: A study of consonant weakening in Spanish-speaking children. In David Eddington (ed.), *Studies in the acquisition of the Hispanic languages: Papers from the 6th conference on the acquisition of Spanish and Portuguese as first and second languages*, 56–65. Somerville, MA: Cascadilla.
- Eremeeva, Victoria & Jane Stuart-Smith. 2003. A sociophonetic investigation of the vowels OUT and BIT in Glaswegian. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona. 1205–1208.
- Foulkes, Paul, Gerard J. Docherty & Dominic Watt. 1999. Tracking the emergence of structured variation: Realisations of (t) by Newcastle children. *Leeds Working Papers in Linguistics & Phonetics* 7. 1–25.
- Foulkes, Paul, Gerard J. Docherty & Dominic Watt. 2005. Phonological variation in child-directed speech. *Language* 81(1). 77–206.
- Guy, Greg & Sally Boyd. 1990. The development of a morphological class. *Language Variation and Change* 2. 1–18.
- Grunwell, Pamela. 1981. *The nature of phonological disability in children*. New York: Academic Press.
- Henry, Alison, Rose MacLaren, John Wilson & Cathy Finlay. 1997. The acquisition of negative concord in non-standard English. In Elizabeth Hughes, Mary Hughes & Annabel Greenhill (eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development*, 269–280. Somerville, MA: Cascadilla Press.
- Hoyle, Susan & Carolyn Temple Adger. 1998. *Kids talk: Strategic language use in later childhood*. New York: Oxford University Press.
- Johnson, Valerie. 2005. Comprehension of third person singular /s/ in AAE-speaking children. *Journal of Language, Speech, and Hearing Services in Schools* 36. 116–124.
- Kerswill, Paul. 1996. Children, adolescents and language change. *Language Variation and Change* 8. 177–202.
- Kerswill, Paul & Ann Williams. 2000. Creating a new town koine: children and language change in Milton Keynes. *Language in Society* 29. 65–115.
- Kerswill, Paul & Ann Williams. 2002. “Salience” as an explanatory factor in language change: evidence from dialect levelling in urban England. In Mari C. Jones & Edith Esch (eds.), *Language change. The interplay of internal, external and extra-linguistic factors*, 81–110. Berlin & New York: Mouton de Gruyter.
- Kerswill, Paul & Ann Williams. 2005. New towns and koineisation: Linguistic and social correlates. *Linguistics* 43(5). 1023–1048.
- Kovac, Ceil, & H. Douglas Adamson. 1981. Variation theory and first language acquisition. In David Sankoff & Henrietta Cedergren (eds.), *Variation omnibus*, 403–410. Edmonton, Alberta: Linguistic Research, Inc.
- Labov, William. 1970. The logic of non-standard English. In Frederick Williams (ed.), *Language and poverty: Perspectives on a theme*. 153–189. Chicago: Markham.

- Labov, William. 1989. The child as linguistic historian. *Language Variation and Change* 1. 85–97.
- Labov, William. 2001. *Principles of linguistic change. Vol. 2: Social factors*. Oxford: Blackwell.
- Leech, Geoffrey. 1987. *Meaning and the English verb*, 2nd edn. London: Longman.
- Macafee, Caroline & Colm O'Baoill. 1997. Why Scots is not a Celtic English. In Hildegard L. C. Tristram (ed.), *The Celtic Englishes*, 245–286. Heidelberg: Carl Winter.
- Macaulay, Ronald K. S. 1991. *Locating dialect in discourse: The language of honest men and bonnie lassies in Ayr*. New York: Oxford University Press.
- Matheson, Catherine & David Matheson. 2000. Languages of Scotland: Culture and the classroom. *Comparative Education* 36(2). 211–221.
- McRae, Sandra. 2000. The demonstrative pronouns in the North-East: An introductory discussion. *Scottish Language* 19. 66–82.
- Miller, Jim. 1993. The grammar of Scottish English. In James Milroy & Lesley Milroy (eds.), *Real English: The grammar of English dialects in the British Isles*, 99–138. New York: Longman.
- Miller, Karen. 2007. *Variable input and the acquisition of plurality in two varieties of Spanish*. East Lansing, MI: Michigan State University dissertation.
- Miller, Karen and Cristina Schmitt. 2009. Syllable-final /s/ lenition and the acquisition of plural morphology in Spanish-speaking children. In John Grinstead (ed.), *Hispanic child languages*, 3–28. Amsterdam & Philadelphia: John Benjamins.
- Murray, James. 1873. *The dialect of the southern counties of Scotland*. London: Philological Society.
- Patterson, Janet Lee. 1992. *The development of sociolinguistic phonological variation patterns for (ing) in young children*. Albuquerque, NM: University of New Mexico dissertation.
- Pinker, Stephen. 1990. Language acquisition. In Daniel N. Osherson & Howard Lasnik (eds.), *An invitation to cognitive science, vol. 1: Language*. Cambridge, MA: MIT Press.
- Purcell, April. 1984. Code-switching Hawaiian style: Children's accommodation along a decreolizing continuum. *International Journal of the Sociology of Language* 46. 71–86.
- Reid, Euan. 1978. Social and stylistic variation in the speech of children: Some evidence from Edinburgh. In Peter Trudgill (ed.), *Sociolinguistic patterns in British English*, 158–171. London: Edward Arnold.
- Renn, Jennifer & J. Michael Terry. 2009. Operationalizing style: Quantifying style shift in the speech of African American adolescents. *American Speech* 84(4). 367–390.
- Roberts, Julie. 1994. *Acquisition of variable rules: (-t,d) deletion and (ing) production in preschool children*. Philadelphia, PA: University of Pennsylvania dissertation.
- Roberts, Julie. 1997a. Acquisition of variable rules: A study of (-t,d) deletion in preschool children. *Journal of Child Language* 24(2). 351–372.
- Roberts, Julie. 1997b. Hitting a moving target: Acquisition of sound changes by Philadelphia children. *Language Variation and Change* 9. 249–266.
- Roberts, Julie. 2002. Child language variation. In Jack Chambers, Natalie Schilling-Estes, & Peter Trudgill (eds.) *Handbook of language variation and change*, 333–348. Oxford: Blackwell.
- Romaine, Suzanne. 1978. Postvocalic /r/ in Scottish English: Sound change in progress? In Peter Trudgill (ed.), *Sociolinguistic patterns in British English*, 144–158. London: Edward Arnold.
- Romaine, Suzanne. 1984. *The language of children and adolescents*. Oxford: Blackwell.
- Smith, Jennifer. 2000a. *Synchrony and diachrony in the evolution of English: Evidence from Scotland*. York: University of York dissertation.

- Smith, Jennifer. 2000b. You Ø na hear o' that kind o' things: Negative *do* in Buckie Scots. *English World-Wide* 21(2). 231–259.
- Smith, Jennifer. 2001. Negative concord in the Old and New World: Evidence from Scotland. *Language Variation and Change* 13(2). 109–134.
- Smith, Jennifer, Mercedes Durham & Liane Fortune. 2007. Community, caregiver and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change* 19(1). 63–99.
- Smith, Jennifer, Mercedes Durham & Liane Fortune. 2009. Universal, dialect-specific pathways of acquisition: Caregivers, children and t/d deletion. *Language Variation and Change* 21(1). 69–95.
- Snow, Catherine. 1994. Introduction. In Clare Gallaway & Brian Richards (eds.), *Input and interaction in language acquisition*, 3–12. Cambridge: Cambridge University Press.
- Stuart-Smith, Jane. 2003. The phonology of modern urban Scots. In John Corbett, Derrick McClure & Jane Stuart-Smith (eds.), *Edinburgh student companion to Scots*, 110–137. Edinburgh: Edinburgh University Press.
- Tagliamonte, Sali. 2006. *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.
- Tagliamonte, Sali, Jennifer Smith & Helen Lawrence. 2005. Disentangling the roots: The legacy of British dialects in cross-variety perspective. In Markku Filppula, Juhani Klemola, Marjatta Palander & Esa Penttilä, (eds.), *Dialects across borders: Selected papers from the 11th international conference on methods in dialectology (Methods XI)*. Joensuu, Finland. August 2002, 87–117. Amsterdam & Philadelphia: John Benjamins.
- Tomasello, Michael. 2003. Constructing a language: A usage-based theory of language acquisition. Cambridge, MA: Harvard University Press.
- Trudgill, Peter. 1996. *Dialects in contact*. Oxford: Blackwell.
- Vosoughi, Soroush. 2010. *Interactions of caregiver speech and early word learning in the Speechome Corpus: Computational explorations*. Cambridge, MA. Massachusetts Institute of Technology MSc thesis.
- Wells, John. 1982. *Accents of English*. Cambridge: Cambridge University Press.
- Westergaard, Marit. 2009. Usage-based vs. rule-based learning: The acquisition of word order in wh-questions on English and Norwegian. *Journal of Child Language* 36. 1023–1051.
- Wolfram, Walt. 1969. *A sociolinguistic description of Detroit Negro Speech* (Urban Language Series 5). Washington, DC: Center for Applied Linguistics.
- Wolfram, Walt. 1991. Re-examining the status of Vernacular Black English: A review article. *Language* 66. 121–133.
- Wolfram, Walt & Ralph Fasold. 1974. *The study of social dialects in American English*. Englewood Cliffs, NJ: Prentice-Hall.
- Wolfram, Walt & Natalie Schilling-Estes. 2006. *American English: Dialects and variation*, 2nd edn. Oxford: Blackwell.
- Yang, Charles. 2002. *Knowledge and variation in natural language*. Oxford: Oxford University Press.
- Youssef, Valerie. 1991. Variation as a feature of language acquisition in the Trinidad context. *Language Variation and Change* 3. 75–101.