University of Pennsylvania Working Papers in Linguistics

Volume 7
Issue 3 *Papers from NWAV 29*

Article 7

1-1-2001

The Emergence of Structured Variation

Paul Foulkes
pfll@york.ac.uk

Dominic J L Watt
djlwl@york.ac.uk

Gerard R. Docherty
g.j.docherty@ncl.ac.uk

The Emergence of Structured Variation

The Emergence of Structured Variation

Paul Foulkes, Gerard J. Docherty and Dominic J.L. Watt

1 Introduction

This paper presents some early findings of an ongoing project, *The Emergence of Structured Variation in the Speech of Tyneside Infants* (*ESV*). This project focuses on the phonetic and phonological development of 40 children ages 2-4, from the city of Newcastle upon Tyne.[1] *ESV* builds upon the findings of an earlier project, which we refer to as *PVC*.[2] *PVC* yielded a detailed record of variation and change among an adult community in Newcastle upon Tyne, showing correlations between various linguistic forms and speech style, gender, class, and age. The project furthermore attempted to integrate methods and theories from variationist sociolinguistics with those of contemporary phonological theory and of acoustic phonetics.

In light of the findings of *PVC*, we are seeking to understand how children acquire complex patterns of structured phonetic variation, including phonetic alternants which are sociolinguistically patterned in the surrounding adult community. In Newcastle English, for example, words of the NURSE lexical set (Wells 1982) may be pronounced with a range of vowel qualities, including [ɔ:] and [ø:]. The former variant is used almost exclusively by males, while the latter is significantly more common in the speech of women (Watt and Milroy 1999, Watt 2000).

Variable features of this kind have received little attention in the literature on child language acquisition. In part this is the result of methodological constraints: most child studies have involved diary-based documentation of the productions of individual children, in contrast with the large samples necessary to locate sociolinguistic variation.

---

[1] *ESV* is funded by the UK Economic and Social Research Council (grant no. R000 237417; 1998-2001). The research team comprises Gerry Docherty and Barbara Dodd from the University of Newcastle, Lesley Milroy from the University of Michigan, and Paul Foulkes and Dominic Watt from the University of York. The fieldworkers are Catherine Ashcroft, Mike Caygill, Marion Hall, Jen Smith, and Debbie Vigil.

[2] *Phonological Variation and Change in Contemporary Spoken British English* (UK ESRC grant no. R000 234892; 1994-1997). For details of this project see Docherty, Foulkes, Milroy, Milroy and Walshaw (1997); Foulkes (1997); Milroy, Milroy, Docherty, Foulkes, and Walshaw (1999); Docherty and Foulkes (1999a); and Watt and Milroy (1999).

Similarly, child-oriented work has played only a minor role within sociolinguistics and phonetics. Local (1983:449) commented that "[r]emarkably little is known about the development and functioning of linguistic variability in the speech of children", although more recent work has provided evidence that variable forms are learned from the beginning of the acquisition process. Local's own work (1978, 1983) investigates the speech patterns of children aged between 4 and 6 from the Newcastle area. He found that variation in vowels and prosodic patterns is structured and changes as phonological rules and representations are mastered. More recently, Hartley's (1992) research with Newcastle 5-year-olds revealed gender-correlated variation in the realization of glottalized forms of voiceless stops (see further Milroy, Milroy, Hartley, and Walshaw 1994). Only a very small number of sociolinguistic studies has been carried out on infants as young as those in the *ESV* project. Wolfram (1989) found evidence for distinctive African-American Vernacular by age 3, while Roberts and Labov (1995) and Roberts (1997a,b) have shown that 3-year-olds in Philadelphia have acquired complex patterns of vowel and consonant realizations, some of which are undergoing change in the adult community. Roberts (1997a) also found statistically significant differences between boys' and girls' usage of particular variants. Roberts and Labov (1995:110) conclude that the 3- to 4-year age range is "a critical period for the acquisition of dialectal norms of the speech community, just as it is for language learning in general".

Studies of children's speech which utilize instrumental phonetic techniques have tended to look for evidence of particular phonological *contrasts* having been acquired. For example, Macken and Barton (1980) investigate the acquisition of the /t/-/d/ contrast in English. However, it has also been shown that children learning different languages display subtle differences in the phonetic forms they use to realize a phonological category. For example, American and Swedish children aged 2;6 differ in place and manner of /t/ production, in accordance with differences found in the speech of American and Swedish adults (Stoel-Gammon, Williams, and Buder 1994). Similar differences were found for vowel duration among the same children (Stoel-Gammon, Buder, and Kehoe 1995), while Scottish children show signs of developing the complex vowel duration patterns known as Scottish Vowel Length Rule from age 4 (see Hewlett, Matthews, and Scobbie 1999, Scobbie, Hewlett, and Turk 1999).

Summarizing what is known about acquisition of variable forms, Chambers (1995:158) comments:

> sociolinguistic competence...develops very early. There are no studies documenting a time gap between the acquisition of linguistic

> competence and the development of sociolinguistic competence. In fact there is no reason to consider them to be different from one another. When children acquire their mother tongues, they evidently acquire the local variants and the norms of their usage too.

Chambers' statement is clearly correct. It is indisputable that children develop their phonological knowledge at least in part via interpretation of the phonetic substance to which they are exposed. As a result, we would expect to find some evidence of sociolinguistically-conditioned variable forms in children's speech production, since they will not know *a priori* that there is any difference between these and other systematic aspects of realization. Distinguishing what is an invariant phonological structure from what is variant and sociolinguistically meaningful is presumably one of the tasks children have to perform in the acquisition process.

There are, however, certain issues embedded within Chambers' summary statement which remain unclear. First, although children may speak with a recognizable accent, showing that they have learned local variants, it is equally obvious that there are always differences between speech produced by children and speech produced by adults. Some of these differences are organic (due to the size and shape of the vocal tract, for instance), but it is possible that other differences might result from children acquiring some variants more quickly or successfully than others. Secondly, sociolinguistic work has demonstrated that accents and dialects, like languages themselves, are not homogeneous entities. Instead, they comprise some features which correlate with the social make-up of speakers, as well as features subject to stylistic variation. In other words, there may be alternatives within what Chambers refers to as the 'norms of usage' of a given dialect, connected, for instance, to the gender or age of the speaker. Where such conditions obtain, how do children interpret the conflicting norms they hear? Do complicated sociolinguistic variables pose a particular problem for the language learner?

These issues of uncertainty form the focus of *ESV*. The specific questions we are therefore addressing include:

- To what extent do young children acquire sociolinguistically-patterned features of their native variety?
- What choices do children make when presented with alternatives, such as gender-correlated variants, in the input they receive?
- At what age do male and female children begin to diverge in their patterns of production?

In this paper, we first present an analysis of the realizations of (t) produced by a subset of the children in the *ESV* study, before addressing some of the theoretical issues which emerge from this preliminary analysis.

2 The *ESV* Project: Methodology

2.1 Sampling

For the *ESV* project as a whole, two bodies of data are being collected in parallel. A cross-sectional study is being undertaken with 40 children—4 boys and 4 girls each at ages 2;0, 2;6, 3;0, 3;6 and 4;0 (± one month), while a subset of the youngest children are also involved in a longitudinal study between the same ages, with recordings being made at bi-monthly intervals. All of the children are drawn from the same neighborhood as the broadly-defined 'working class' Newcastle cohort of the *PVC* study. All subjects are only children or the eldest child, in order to control for the impact of communication with siblings. The parents are in all cases the primary care-givers, and in most families the mother plays the larger role. The children are all monolinguals, and have no known speech or hearing disorders.

The research reported in this paper focuses on a subset of 26 children from the cross-sectional study, with analysis of the full cohort in progress.

2.2 Recording and Analysis

The children are recorded for a total of approximately 45 minutes, in the contexts of (i) free play sessions with the mother and a fieldworker, and (ii) carrying out a toy- and picture-based word elicitation procedure. The mother is also recorded reading a word-list. Recordings are made using Trantec lapel radio microphones and a Sony TCD-D10 Pro II DAT recorder. Both auditory and acoustic analysis of variables are subsequently undertaken, the latter using Sensimetrics *SpeechStation 2*. For the analysis of (t) discussed in the present paper we have concentrated on using acoustic analysis to identify the features produced by the children, since our analysis of adult speech has shown that degrees of variability in (t) realization are often more complex than auditory analysis permits (Docherty and Foulkes 1999a,b). The entire tape recording for each child has been analyzed for this purpose.

2.3 *ESV* Variables

Since our aim in this project is to look at how variable features emerge in the acquisition process, our approach differs from that taken in the majority of

child-centred work in that we are not principally concerned with the emergence of phonemic contrasts. Analysis in *ESV* centres on a range of consonant and vowel variables which were investigated in the adult study, *PVC*. These include realizations of (t), as discussed in the present paper; glottalization of (p) and (k); and the quality of the vowels in the NURSE, FACE, and GOAT lexical sets. Comparisons are drawn between the findings for the *ESV* children and the *PVC* adults (who were sampled from the same Newcastle community), as well as between the *ESV* children and their mothers. Another ongoing branch of the *ESV* work compares the phonological/phonetic features of child-directed speech with those of adult-to-adult speech (Docherty, Foulkes, Parsons, Thompson, Tillotson, and Watt 2000).

3 Variable (t)

The variable (t) is particularly complex in adult speech (see especially Docherty *et al.* 1997, and Docherty and Foulkes 1999a). First, adults use different variants in word-initial, intersonorant, and word-final pre-pausal positions. Second, the intersonorant and pre-pausal variants are subject to sociolinguistic patterning, with gender appearing to be a particularly important factor. Third, lexical constraints are found, such that certain variants occur only in a restricted set of words. Specific details are given in sections 3.1 to 3.4 below.

The variable (t) has been analyzed in four contexts, as shown in Table 1.

context	**abbreviation**	**examples**
word-initial	WI	*toy, Teletubby*
word-medial intersonorant	WMIS	*water, bottle*
word-final intersonorant	WFIS	*get off, sat on*
word-final pre-pausal	WFPP	*cat #, parrot #*

Table 1. Variable (t): contexts of analysis, abbreviations, examples.

Three of the four contexts (WMIS, WFIS, and WFPP) are known from the *PVC* study to be the locus of socially-sensitive variation in the adult community. The fourth context, WI, was not investigated in the *PVC* study, but is included in *ESV* so that we may take stock of the entire range of (t) variants produced by the children.

The following four sections (3.1 to 3.4) focus on these contexts in turn. We first outline the patterns found for adult speakers, drawing attention to the phonetic forms used, and any sociolinguistic and/or lexical patterning apparent. The adult data are drawn from the *PVC* study for the WMIS, WFIS, and WFPP contexts. We then describe the findings from examination of the children's data. Section 3.5 offers a summary of the findings.

3.1 Word-initial (t)

Most varieties of English have voiceless stops that are aspirated when in a stressed word-initial position. That is, the release of the oral closure is followed by a delay in voice onset time (VOT) of at least 25 milliseconds. Studies of English-speaking adults have found that the VOT of initial [t^{h}-] tends to fall in the 30-105 ms range, with mean values around 65-70 ms (e.g. Lisker and Abramson 1964; Kewley-Port and Preston 1974). VOT is also one of the few topics to have been investigated by instrumental phonetic means in acquisition studies. Previous work has shown that 2-year-old English-speaking children can usually contrast initial voiced and voiceless stops by means of different VOT durations (Macken and Barton 1980). Also by age 2 years, children produce voiceless stops with long lag VOT, but with a longer average duration and greater variability than do adults. Adult-like consistency is achieved at around 5 years (e.g. Kewley-Port and Preston 1974, Gilbert 1977, Smith 1978).

We will deal with the *ESV* findings very briefly here, since they are entirely in line with the findings of these previous studies. From the *ESV* recordings we extracted all instances of word-initial /tV-/, and measured the VOT duration from the release burst of the plosive to the onset of periodic vocal fold vibration.

The number of tokens per child ranged from 5 to 57, with a mean of 26. All subjects produced 95% or more of their tokens with long lag VOT. The lowest average VOT duration was 39.09 ms and the highest 113.2 ms. Fourteen children had VOT averages greater than 70 ms, which is the value typically cited for English-speaking adults (Lisker and Abramson 1964). The average VOT duration for all 26 children was 79.95 ms, which is closely comparable with child data from Smith (1978) and others. The high average standard deviation of 44.47 ms shows that there is also a good deal of variability in the children's speech patterns. However, the data as a whole demonstrate that even the youngest children had mastered appropriate performance strategies for [t^{h}-].

3.2 Word-medial Intersonorant (t)

Acoustic analysis of the *PVC* adults showed that 70% of tokens in the WMIS context (e.g. *water*) were produced with unbroken voicing, and 93% included an interval of creaky phonation. Only 3% of tokens were fully voiceless (Docherty and Foulkes 1999a,b). For this reason we use the transcription [d̰] to represent the default variant of (t) in this context, and we

assume that children learning this dialect must target both voiced and creaky realizations in order to match those of the adults.

Our acoustic analysis of the children's speech has therefore focused on identifying the presence of full voicing and creak. A token is defined as fully voiced if periodic noise is apparent throughout the target segment. Creak is defined as being present if it can be observed either during the target segment, or within the syllable nucleus either side of the stop (allowing for the fact that adults also show variation in temporal co-ordination of the appropriate gestures; Docherty and Foulkes 1999a,b).

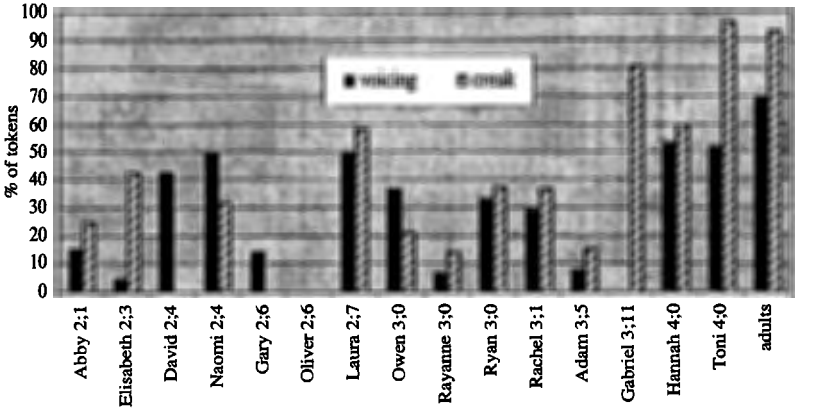

Figure 1. Proportion of WMIS tokens with full voicing and creak.

The findings of the acoustic analysis are displayed in Figure 1. This graph shows the proportion of tokens containing full voicing and creak, as produced by the 15 children who produced at least 12 tokens in this context. The total number of tokens is 280, with a mean of 20. The children are arranged in order of increasing age, from left to right. Data from the *PVC* young working class adults are shown for comparison.

Figure 1 suggests there is evidence that the children are reproducing the key features found in the adult data, but to varying degrees. Creak and full voicing occur in about a third of tokens in the children's speech, whereas both are overwhelmingly present in adult speech. There may also be signs of age-grading: The two eldest children display patterns which are most similar to those of adults, as does Laura, who is linguistically advanced for her age.

In general, it is clear that the data for WMIS position show less of a match between children's and adults' patterns than those for WI position.

However, our ongoing research into the phonological patterns of child-directed speech complicates our interpretation of these data. Although [d̰] is the norm in adult-to-adult speech, many of the mothers use a substantial proportion of a more standard-like form [t^{h}] when addressing children (Docherty *et al.*, 2000). In light of this finding, the commonest form heard in the ambient language may differ from child to child. We intend to assess the correlation between each mother-child pairing before drawing firm conclusions about variant learning in this context. What remains clear, however, is that almost all children produce a significant number of WMIS tokens (involving voicing and/or creak) which differ from those used in WI position (which involve long lag VOT), showing that context-dependent forms have to a reasonable extent been differentiated by children in the learning process.

3.3 Word-final Intersonorant (t)

As in WMIS position, word-final (t) may be realized as [d̰] when a vowel follows. In the *PVC* data, these variants accounted for 34% of the WFIS tokens produced by young working class (WC) speakers, with no significant difference between the sexes (Docherty *et al.* 1997:293 [table 4]). There are, however, other variants which are possible in this context, including [ɹ], [d], [ʔ], [ɾ], and [t]. The variant [t] accounted for only around 5% of the young WC adult sample, which enables us to generalize that 95% of tokens were voiced (since the forms transcribed [ʔ], like [d̰], are usually characterized by unbroken voicing; see Docherty and Foulkes 1999a,b).

The variant [ɹ] introduces two complexities. First, it occurs only in a restricted set of lexical items, mainly common monosyllabic verbs (*get, got, put*) and non-lexical words (*that, not, but, what*). Second, it is significantly more frequent in the speech of females than that of males. In the young WC data, for example, [ɹ] occurred in 21% of female tokens compared with just 3% in the male tokens (Docherty *et al.* 1997:293 [table 4]).

In the *ESV* data the principal features sought in the acoustic analysis are therefore presence of full voicing and presence of creak. Given their auditory salience, [ɹ]-like variants were primarily identified by auditory rather than acoustic analysis, with particular attention being paid to the lexical items in which they occur.

The acoustic findings are shown in Figure 2. Data are from 15 children, with a minimum of 8 tokens per child, a mean of 19, and a total of 287.

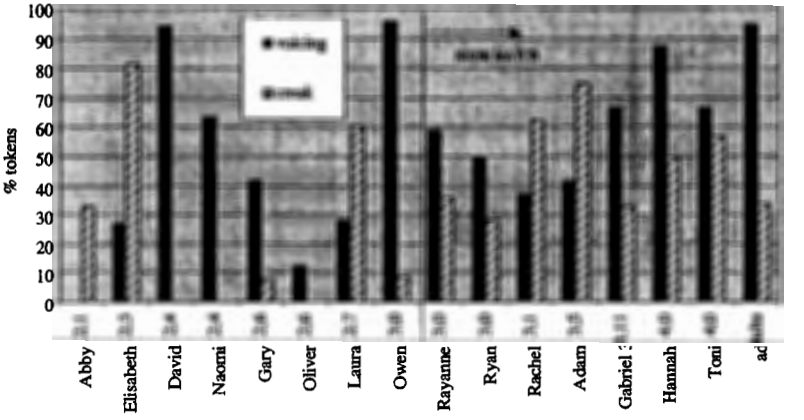

Figure 2. Proportion of WFIS tokens with full voicing and creak.

Comparing Figures 1 and 2, it is striking that the overall match between children and adults appears closer in WFIS position (Figure 2), especially for the older children, shown towards the right of the graph. Perhaps most striking, however, is that the children produce over twice as many voiced variants of (t) in WFIS position (overall 56% in Figure 2) than in WMIS position (30% in Figure 1). The children therefore show different realizational patterns in accordance with the different patterns produced by the adults.

The [ɹ] variant occurs 19 times from a possible maximum of 175 (the remaining 112 WFIS tokens involve lexical items which would not accommodate [ɹ] in adult speech). In all cases the realization is actually more like a labiodental or bilabial approximant, [ʋ], which is well known to be common in children's speech but is also becoming more widespread as a mature accent feature in many British varieties (Foulkes and Docherty 2000). Although [ɹ] is significantly associated with female speech among adults, the majority of tokens among children are produced by boys (14/89, or 16%; the girls produce 5/86, or 6%).[3] All 19 cases of [ɹ] occur in appropriate lexical items (*got, get, what* etc), showing that the children who do use it have acquired knowledge of the lexical selectivity of the variant. This finding is further confirmed by the absence of [ɹ] forms in any other context.

[3] Statistical analysis will be carried out once the full cohort has been analyzed.

3.4 Word-final Pre-pausal (t)

In contrast with the contexts described in 3.2 and 3.3 (as well as in contrast with most other British accents), pre-pausal (t) is only rarely realized as any sort of glottal or laryngealized/creaky form in Newcastle English. More common variants in this context include [tʰ] and, for some speakers, pre-aspirated stops, [ʰt]. In acoustic terms the latter may appear either as a period of high-frequency frication *before* the voiceless stop gap, or as a breathy continuation of a preceding vowel (see further Laver 1994:356-7; Docherty and Foulkes 1999a). The rarer creaky/glottal forms generally occur in a restricted set of words. Coincidentally, this set is very similar to that which accommodates the [ɪ] variant in WFIS context (section 3.3) such as *what, it, that, got*. Several of these occur as part of sentence tags such as *and that* and *isn't it?* (Docherty *et al.* 1997). The acoustic correlates of these variants include creaky voice (as in section 3.2), which may be accompanied by patterns indicative of a glottal plosive articulation.

Both the pre-aspirated and glottal variants are socially sensitive within the local adult community. Pre-aspirated variants are significantly associated with working class female speech amongst young adults, and are virtually absent in the 45-65 age group (Docherty and Foulkes 1999a). Analysis of the *PVC* data revealed 63% of WFPP tokens produced by young WC females to be pre-aspirated, against 13% for the males. Similarly, the glottal forms in pre-pausal position are also almost entirely exclusive to the speech of young working class women (Docherty *et al.* 1997; Docherty and Foulkes 1999a).

Analysis of the children's data has again concentrated on establishing the presence of key acoustic features, namely pre-aspiration and creak. A token is classified as pre-aspirated if it contains a period of distinct fricative energy prior to the stop closure, or a breathy voiced offset to the preceding vowel. To enable comparison with other contexts we also assessed the presence of full voicing.

Figure 3 presents the proportion of tokens classified as pre-aspirated for all 26 children so far analyzed, along with overall scores for the young WC women and men from the *PVC* study. The children produced a minimum of 12 tokens of (t) in this context, with a mean of 38 and a total of 974.

The first point to note is the high incidence of pre-aspiration, which, as indicated above, is an innovative feature of the local accent, being strongly associated with young females. Across all 26 children, 46% of tokens display pre-aspiration (cf. 63% by women and 12% by men). The gross finding therefore points to a tendency among the children to adopt a pattern predominant in female speech where one is available, a finding which therefore supports the predictions of Labov (1990) and the findings of Roberts (1997a,b). Taken as groups, the boys and the girls show no clear

differences in usage, displaying the pre-aspirated pattern in roughly equal measure (49% versus 42%).

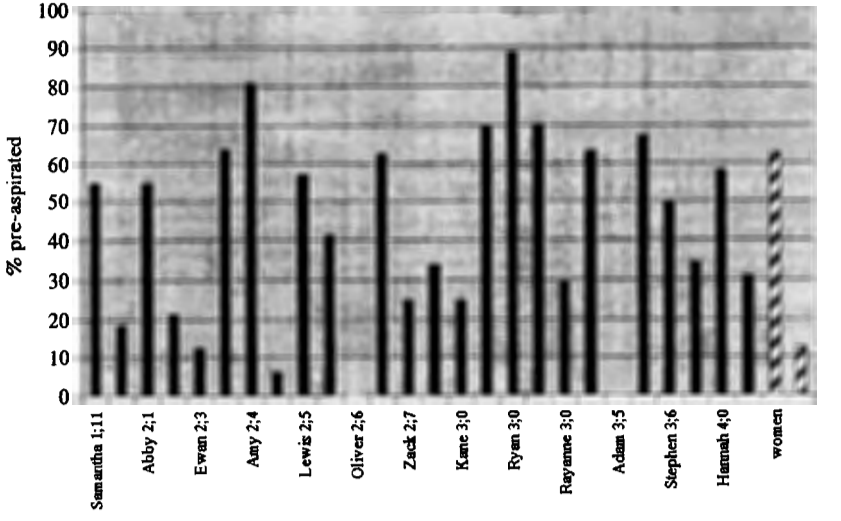

Figure 3. Proportion of WFPP tokens with pre-aspiration.

It is also apparent that there is a relatively wide range of variation among individuals in Figure 3. Preliminary analysis comparing the use of pre-aspiration in 17 mother-child pairs shows a weak correlation for the whole cohort (r = .368, just missing significance at the 10% level). However, a very strongly significant correlation is found when three outliers—mothers with low pre-aspiration scores whose children score highly—are removed from the analysis (r = .663, p < .005; Docherty *et al.* 2000). This clearly points to a close connection between forms used by mothers and those initially acquired by their children, although we must wait for the analysis to be completed before we draw firm conclusions.

With regard to the other acoustic features analyzed, voicing is overwhelmingly absent, as we would predict. Recalling that 56% of word-final tokens were fully voiced when followed by a vowel (section 3.3), we can conclude that the children are in general sensitive to this phonologically-governed process of the dialect. Creaky or glottal tokens account for 18% of the total, substantially less than in the WMIS and WFIS contexts. We also find that around two thirds of the tokens involving creak occur in the

restricted lexical set where glottals may be used by adults, and particularly by young WC women.

3.5 Summary of Findings

To recapitulate the patterns discussed in sections 3.1 to 3.4, we find that positional differentiation of phonetic forms is well established. The children produce voiceless aspirated stops in word-initial position; a high proportion of voiceless pre-aspirated stops pre-pausally; and a range of variants in medial positions which are not found initially or pre-pausally, including a significant number of the salient local variants found in the adult community.

The children therefore appear to have made good progress in acquiring different variants for the different phonological contexts. The match to adult forms is in some cases less successful than in others, notably in the WMIS and WFIS contexts. There is some evidence that older children produce a closer match to the adult patterns than do the younger children.

Where sociolinguistic differences exist in the adult community, resulting in a degree of choice of target for the children, performance is again variable. Pre-aspiration of pre-pausal (t) and the WFIS [ɪ] variant are predominant in young women's speech, and are also common for the children. The children display complete mastery of the lexical restriction on [ɪ]. At this stage of the project it is too early to draw any firm conclusions with regard to the boys' versus the girls' performance, but no particular differences are apparent in this sample of children.

4 Discussion

In this final part of the paper we offer a brief discussion of some of the theoretical implications of our findings.

Our data show, first of all, that it is vital to take account of patterns of variability in the adult community, as well as in child-directed speech, in order to make a full and accurate assessment of the targets children aim for in acquisition. We therefore support Edwards (2000:245) in calling for a more sophisticated methodology in phonological acquisition studies. Edwards points out:

> we cannot consider acquisition of a phoneme category independently of the segmental and prosodic contexts in which that phoneme is produced. We can't simply ask "Can the child produce /k/?" Rather, we must ask whether a child can produce /k/ word-

initially and word-finally, in clusters, before vowels, after back vowels, in stressed onsets, in unstressed codas, and so on.

To this list we would also add that children must learn to reproduce sociolinguistically appropriate variants.

Failure to take account of such factors may lead to the view that variability in children's performance is the result of a high degree of inconsistency. Variability itself is widely recognized as one of the most obvious characteristics of children's speech. However, linguists researching children's speech have tended to try to reduce this variability, concentrating on characterizing the patterns of *simplification* between child and adult forms (see further MacNeilage 1980, 1997). Variability has therefore often been attributed to imperfect learning, or to universal effects of immature development, or to the actions of default phonological processes. For instance, glottalization and deletion of /t/ in coda position has been analyzed as the result of a phonological rule of simplification (see Locke 1983:230 ff.).

However, the variability in (t) realizations among our children is clearly appropriate given the input they receive. This further implies that the variability is mainly the result of extremely sophisticated and detailed learning rather than the product of error or imperfection. Consider, in support of this conclusion, the articulatory activity which the children control to produce their range of (t) realizations. In initial position, a stop with long lag VOT involves first of all oral closure with glottal opening, followed by a glottal closing gesture to initiate voicing. Crucially, the glottal closing must be timed to follow the release of the oral stop by at least 25 ms. In WMIS position the variant [d̰] requires an oral gesture similar to that for [tʰ] in WI position, but with voicing throughout, and coupled with a laryngeal adjustment to produce a brief period of creaky phonation. In WFIS position this same variant may occur, as may various other voiced phones including [ɹ], which involves a difference in the oral constriction(s) involved. Finally, in WFPP position, variants are usually voiceless, but the co-ordination of oral and laryngeal gestures is once again different from that in initial position: the voicing must be switched off *before* the oral closure is made. The (t) variants therefore involve subtle and highly complex differences in the co-ordination of oral and laryngeal gestures.

These findings are furthermore consistent with models of phonological acquisition which place a degree of emphasis on 'bottom-up' learning (e.g. Vihman and Velleman 2000). That is, the ambient language plays a significant role in driving the development of phonological structures.

Support for a heavily 'bottom-up' acquisitional process entails predictions with regard to the form early lexical representations take, and

also to the specific phonological structures (phonemes, features, syllables etc.) which these representations may comprise. Our data are consistent with views that early lexical representations are detailed and holistic in nature, and might contain relatively little in the way of abstract phonological information (e.g. Studdert-Kennedy 1983, Pisoni 1993). Multiple trace models of representation (e.g. Hintzman 1986, Jusczyk 1997) further allow different representations of single lexical items to be stored. Some normalization may take place as part of the initial learning process, but a substantial degree of detail may remain. As a result, lexical representations are not conceived of as minimalistic structures such as phoneme strings.

In the case of Newcastle (t), for example, this might mean that many different exemplars of words such as *what* may be learned, each memorized to include details about the speaker and communicative situation, as well as fine-grained phonetic material such as the physical form produced in the final portion of the word. Meta-linguistic analysis may develop after an initial lexical store has been established, resulting in a more abstract analysis of the phonological components of words. It may therefore take some time, for example, for a Newcastle child to establish an abstract phonological component to link the final stages of words such as *what* with words such as *cat*. Establishing a link between these word components and those of, say, *teddy* and *toy* may require further analysis. At the same time, the multiple instances of a particular category are thought to converge on a smaller set of relevant categories (Jusczyk 1992, 1993). That is, the multiple traces of a word may be reduced to a single, more abstract, lexical entry.

If we follow models of this type, we can predict that the details encoded at the holistic stage of representation will include features which have sociolinguistic relevance among the adult community. There should also be a bias towards forms that are heard most often, which for many children means the forms produced by younger women. The learning of detailed variant forms may therefore be in progress well before lexical representations are reorganized to reflect abstract phonological structure. If we are right in this conclusion, we might be persuaded that Chambers (1995) is actually wrong to claim that linguistic and sociolinguistic competence are learned together: aspects of sociolinguistic detail in fact *precede*, and contribute to, the development of phonological knowledge.

This suggests, contrary to widespread assumptions, that variability may play a positive role in acquisition. Variability in the speech signal has typically been characterized as disfunctional (see Pisoni 1997 for a review). For example, it has been shown in listening experiments that lexical processing is faster when a single speaker's voice is heard than when several different voices are involved (Pisoni 1997, Mullennix 1997). Signal

variability has also been characterized as disfunctional with respect to acquisition. It has been suggested that variant forms may delay or hinder the development of adult-like phonological categories such as phonemes (see e.g. Locke 1983:202-3). These categorial units must be extracted from their encoding within the phonetic medium. It therefore follows logically that the process of extraction will involve a good deal more cognitive effort when a unit is encoded in many different ways, e.g. in a range of different phonetic variants.

However, recent work in adult speech perception has identified a more positive role for variation. It has been shown that the task of learning phonological categories in a second language is improved by exposure to multiple voices. For instance, Lively, Logan, and Pisoni (1993) show that Japanese learners of English are better able to distinguish /l/ from /r/ after hearing input from several speakers rather than a single speaker. They are also helped in reproducing the distinction in their own speech, and at generalizing their perceptual ability to new voices. Lively *et al.* conclude that variability aids the learners in constructing robust phonological categories. Our findings suggest that first language acquirers may benefit from variability in similar ways.

5 Conclusion

The data presented here are preliminary, and our conclusions necessarily somewhat tentative. Ongoing work focuses on (t) productions of the remaining children, including a longitudinal study of the speech of a subset of the group. Numerous other variables are also being studied, which will enable us to draw firmer conclusions on the issues raised in this paper. However, certain clear points emerge. Our findings suggest first of all that it is not straightforward to separate acquisition of contrastive phonological units from acquisition of variable features which make up a child's developing sociolinguistic identity. Context- and accent-specific variants are present very early, although some seem to develop more slowly than others.

There may be evidence to support Labov's (1990) predictions of bias towards female patterns in the input, which has implications for the transmission of sound changes. The apparent closeness of the children's patterns to those of their mothers furthermore raises important questions concerning the actual targets which the children aim for in acquisition. Patterns which are particular to the mothers' speech, including features which are undergoing change in the local accent, appear to be acquired readily by the children. Acquisition studies which do not take account of

accent-specific targets may therefore be inappropriately assessing the child's task in acquisition.

Finally, the initial findings of the *ESV* study support a growing body of work in which variation is viewed not as a design fault of language, but as an integral, functional aspect of human communication in its social context.

References

Chambers, Jack K. 1995. *Sociolinguistic Theory*. Oxford: Blackwell.

Docherty, Gerard J. and Paul Foulkes. 1999a. Derby and Newcastle: instrumental phonetics and variationist studies. In *Urban Voices*, ed. Paul Foulkes and Gerard J. Docherty, 47-71. London: Arnold.

Docherty, Gerard J. and Paul Foulkes. 1999b. Sociophonetic variation in 'glottals' in Newcastle English. *Proceedings of the XIVth International Congress of Phonetic Sciences*, 1037-1040. University of California, Berkeley.

Docherty, Gerard J., Paul Foulkes, James Milroy, Lesley Milroy and David Walshaw. 1997. Descriptive adequacy in phonology: a variationist perspective. *Journal of Linguistics* 33:275-310.

Docherty, Gerard J., Paul Foulkes, Elizabeth Parsons, Jaclyn Thompson, Jennifer Tillotson and Dominic J.L. Watt. 2000. Phonological variation in child-directed speech. Paper presented at NWAVE 29, Michigan State University, October 6.

Edwards, Jan. 2000. Commentary. Lexical representations in acquisition. In *Papers in Laboratory Phonology V. Acquisition and the Lexicon*, ed. Michael B. Broe and Janet B. Pierrehumbert, 240-249. Cambridge: Cambridge University Press.

Foulkes, Paul. 1997. Rule inversion in a British English dialect - a sociolinguistic investigation of [r]-sandhi in Newcastle upon Tyne. *University of Pennsylvania Working Papers in Linguistics* 4(1) - *A Selection of Papers from NWAVE 25*. 259-270.

Foulkes, Paul and Gerard J. Docherty. 2000. Another chapter in the story of /r/: 'labiodental' variants in British English. *Journal of Sociolinguistics* 4:30-59.

Gilbert, J.H.V. 1977. A voice onset analysis of apical stop production in three-year-olds. *Journal of Child Language* 4:103-110.

Hartley, Sue. 1992. A study of the effect of sex and age on glottalisation patterns in the speech of Tyneside schoolchildren. Undergraduate dissertation, University of Newcastle upon Tyne.

Hewlett, Nigel, Ben Matthews and James M. Scobbie. 1999. Vowel duration in Scottish English speaking children. *Proceedings of the XIVth International Congress of Phonetic Sciences*, 2157-2160. University of California, Berkeley.

Hintzman, D.L. 1986. Schema abstraction in a multiple-trace memory model. *Psychological Review* 93:411-423.

Jusczyk, Peter. 1992. Developing phonological categories from the speech signal. In *Phonological Development: Models, Research, Implications*, ed. Charles A. Ferguson, Lise Menn and Carol Stoel-Gammon, 17-64. Timonium, MD: York Press.

Jusczyk, Peter. 1993. From general to language-specific capacities: the WRAPSA Model of how speech perception develops. *Journal of Phonetics* 21:3-28.

Jusczyk, Peter. 1997. *The Discovery of Spoken Language*. Cambridge, Mass.: MIT Press.

Kewley-Port, D. and M.S. Preston. 1974. Early apical stop production: a voice onset time analysis. *Journal of Child Language* 2:195-210.

Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2:205-254.

Laver, John. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

Lisker, Leigh and Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20:384-422.

Lively, S.E., J.S. Logan and D.B. Pisoni. 1993. Training Japanese listeners to identify English /r/ and /l/: the role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America* 94:1242-1255.

Local, John K. 1978. Studies towards a description of the development and functioning of children's awareness of linguistic variability. Doctoral dissertation, University of Newcastle upon Tyne.

Local, John K. 1983. How many vowels in a vowel? *Journal of Child Language* 10:449-453.

Locke, John L. 1983. *Phonological Acquisition and Change*. New York: Academic Press.

Macken, M.A. and D. Barton. 1980. A longitudinal study of the acquisition of the voicing contrast in American-English word-initial stops, as measured by voice onset time. *Journal of Child Language* 7:41-74.

MacNeilage, Peter F. 1980. The control of speech production. In *Child Phonology. Volume 1: Production*, ed. G.H. Yeni-Komshian, J.F. Kavanagh and C.A. Ferguson, 9-21. New York: Academic Press.

MacNeilage, Peter F. 1997. Acquisition of speech. In *The Handbook of Phonetic Sciences*, ed. W.J. Hardcastle and J. Laver, 301-332. Oxford: Blackwell.

Milroy, James, Lesley Milroy, Sue Hartley and David Walshaw. 1994. Glottal stops and Tyneside glottalization: competing patterns of variation and change in British English. *Language Variation and Change* 6:327-357.

Milroy, Lesley, James Milroy, Gerard J. Docherty, Paul Foulkes and David Walshaw. 1999. Phonological variation and change in contemporary English: evidence from Newcastle upon Tyne and Derby. In *Cuadernos de Filología Inglesa 8. Variation and Linguistic Change in English*, ed. J.C. Condé Silvestre and J.M. Hernández-Campoy, 35-46. University of Murcía.

Mullennix, J.W. 1997. On the nature of perceptual adjustments to voice. In *Talker Variability in Speech Processing*, ed. K. Johnson and J.W. Mullennix, 67-84. San Diego: Academic Press.

Pisoni, David B. 1993. Long-term memory in speech perception: some new findings on talker variability, speaking rate and perceptual learning. *Speech Communication* 13:109-125.

Pisoni, David B. 1997. Some thoughts on 'normalization' in speech perception. In *Talker Variability in Speech Processing*, ed. K. Johnson and J.W. Mullennix, 9-32. San Diego: Academic Press.

Roberts, Julie. 1997a. Acquisition of variable rules: a study of (-t, d) deletion in preschool children. *Journal of Child Language* 24:351-372.

Roberts, Julie. 1997b. Hitting a moving target: acquisition of sound change in progess by Philadelphia children. *Language Variation and Change* 9:249-266.

Roberts, Julie and William Labov. 1995. Learning to talk Philadelphian. *Language Variation and Change* 7:101-112.

Scobbie, James M., Nigel Hewlett and Alice Turk. 1999. Standard English in Edinburgh and Glasgow: the Scottish vowel length rule revealed. In *Urban Voices*, ed. Paul Foulkes and Gerard J. Docherty, 230-245. London: Arnold.

Smith, B.L. 1978. Temporal aspects of English speech production: a developmental perspective. *Journal of Phonetics* 6:37-67.

Stoel-Gammon, Carol, Eugene Buder and Margaret Kehoe. 1995. Acquisition of vowel duration: a comparison of Swedish and English. *Proceedings of the XIIIth International Congress of Phonetic Sciences* 4:30-36. University of Stockholm.

Stoel-Gammon, Carol, K. Williams and Eugene Buder. 1994. Cross-language differences in phonological acquisition: Swedish and American /t/. *Phonetica* 51:146-158.

Studdert-Kennedy, Michael. 1983. On learning to speak. *Human Neurobiology* 2:191-195.

Vihman, Marilyn M. and Shelley Velleman. 2000. Phonetics and the origin of phonology. In *Phonological Knowledge: Conceptual and Empirical Issues*, ed. Noël Burton-Roberts, Philip Carr and Gerard J. Docherty, 305-339. Oxford: Oxford University Press.

Watt, Dominic J.L. 2000. Phonetic parallels between the close-mid vowels of Tyneside English: are they internally or externally motivated? *Language Variation and Change* 12:69-101.

Watt, Dominic J.L. and Lesley Milroy. 1999. Patterns of variation and change in three Newcastle vowels: is this dialect levelling? In *Urban Voices*, ed. Paul Foulkes and Gerard J. Docherty, 25-46. London: Arnold.

Wells, John C. 1982. *Accents of English* (3 vols.). Cambridge: Cambridge University Press.

Wolfram, Walt. 1989. Structural variability in phonological development: final nasals in Vernacular Black English. In *Language Change and Variation*, ed. Ralph Fasold and Deborah Schiffrin, 301-332. Amsterdam: John Benjamins.

(Paul Foulkes, Dominic Watt)
Department of Language and Linguistic Science
University of York
York
YO10 5DD
UK
pfl1@york.ac.uk
djlw1@york.ac.uk

(Gerry Docherty)
Department of Speech
University of Newcastle
Newcastle upon Tyne
NE1 7RU
UK
g.j.docherty@ncl.ac.uk