# TRACKING THE EMERGENCE OF STRUCTURED VARIATION – REALISATIONS OF (t) BY NEWCASTLE CHILDREN

Paul FOULKES, Gerry DOCHERTY & Dominic WATT

**Abstract**

This paper describes an investigation of the speech of children aged 2 to 4 from Newcastle upon Tyne. Our aim is to understand how variant phonetic patterns come to be acquired. These patterns include both phonologically governed alternations, such as the aspirated allophone of word-initial (t), and also sociolinguistically correlated variants. A baseline for this study is provided by a previous project focusing on Newcastle adults, which revealed variation and change in the pronunciation of several consonants and vowels. We concentrate here on 10 children's productions of (t), a particularly complex variable in adult speech. The children demonstrate a sophisticated mastery of many aspects of the adult patterns, producing qualitatively different phonetic variants in appropriate phonological contexts. The acoustic qualities of the allophones in general closely resemble those of adults. Where gender-based differences occur in the local adult community, both boys and girls adhere more closely to the patterns typical of women. This suggests that at this stage of development children are most influenced by the phonological/phonetic patterns of their mothers. We conclude that it is problematic to view acquisition of language-specific phonological units as separate from the accent-specific variation which is important in the construction of a sociolinguistic identity. We also suggest that structured variation in the adult input, although usually characterised as dysfunctional in child-centred research, may in fact serve a positive function in the acquisition process. Specifically, children may exploit recurrent patterns in the input signal to help them locate the phonological components of words and achieve the transition from a system of holistic lexical representation to one involving more abstract categorial units.

## 1. Introduction

This paper presents preliminary findings from an ongoing project entitled *The Emergence of Structured Variation in the Speech of Tyneside Infants* (henceforth *ESV*). The *ESV* project focuses on the speech of 40 children aged between 2 and 4 from Newcastle upon Tyne.[1]

In this project we address a series of questions that emerged from another recently completed project (referred to as *PVC*), which investigated sociolinguistic patterns in an adult community of Newcastle upon Tyne.[2] In light of the findings of

[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 Leeds. We wish to thank the fieldworkers Catherine Ashcroft, Mike Caygill, Marion Hall, Jen Smith and Debbie Vigil. We are also grateful for comments and advice we have received from Ghada Khattab, Nicole Whitworth, and from colleagues who attended the Child Phonology Conference (University of Wales, Bangor, July 1999) and the International Congress of Phonetic Sciences (San Francisco, August 1999), at which versions of this paper were presented.

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

*PVC*, we are seeking to understand how complex patterns of structured phonetic variation come to be acquired in the process of language learning and established in speech performance. Our aim is to track the path taken by children in learning the full inventory of allophonic alternants characteristic of the accent to which they are exposed.

This inventory comprises what might be described as two different types of alternant. On the one hand we are investigating those alternants which linguists usually classify as *phonologically governed*. That is, the appearance of a particular phonetic form can be analysed as linked to the phonological structure of a word or phrase. Examples from English include /t/ being articulated as aspirated [$t^h$-] when in word-initial position, or (in many dialects) as a glottal stop when in word-final pre-consonantal position. Learning such alternants is clearly part of the task of acquiring the phonology of a language, and many previous acquisition studies have analysed variables of this kind. The acquisition of aspirated word-initial stops by children learning English, for example, has been documented by Kewley-Port & Preston (1974), Gilbert (1977), and Macken & Barton (1980), among others.

On the other hand, we are also interested in those alternants which would typically be defined as *sociolinguistically governed*. In Newcastle English, for example, words of the NURSE lexical set may be pronounced with a range of vowel qualities, including [ɔ ] and [ø ]. The former variant is used almost exclusively by men, while the latter is significantly more common in female speech (Watt & Milroy 1999). Unlike phonologically governed variables, however, this type of variable has received little attention in the literature on the early stages of acquisition (see section 2 below for a discussion of the few previous studies along these lines). Nevertheless, it is clear that children do acquire an accent of some sort from the earliest age, and presumably this 'accent' may contain elements which have a sociolinguistic value in the adult community.

Our working hypothesis on this issue is as follows: if children develop their phonological abilities primarily through interpretation of the phonetic substance to which they are exposed, we expect to see some influence of sociolinguistic variants, since the child will not *a priori* know that there is any difference between these and other systematic aspects of realisation (including phonologically governed alternants). Distinguishing the role(s) and the significance of the two types of variant is presumably one task children have to perform in the acquisitional process. Once again, however, traditions in previous acquisition work mean we have little understanding of how children might achieve this task. As MacNeilage (1997: 326) points out, the bulk of work on phonological acquisition has concentrated on characterising the patterns of simplification between child and adult forms. In earlier work, MacNeilage (1980) has attributed this to the dominant theoretical tenets of acquisition work, couched largely within the structuralist/generative model. In particular, MacNeilage points to the enduring influence of Jakobson, who viewed the phonetics of child speech as largely irrelevant for an understanding of the acquisition of phonological contrasts. We must furthermore acknowledge the effects of methodological constraints on addressing variation in child language: 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.

Summarising the canon of work in 1983, Local commented that '[r]emarkably little is known about the development and functioning of linguistic variability in the speech of children' (Local 1983: 449). It seems that little progress has been made in

this direction over the last 16 years. Our aim in the *ESV* project is to bridge this gap, by addressing questions such as:

- to what extent do young children acquire sociolinguistically-relevant features of their native variety?
- how do these features relate to the children's mastery of phonologically-governed, context-dependent allophones?
- to what extent do the features adopted by children match those of the adult community?
- what choices do children make when presented with gender-correlated allophones 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 offer a brief review of previous work which has addressed the acquisition of variant forms. Then, in the main body of the paper, we turn to an analysis of the realisations of (t) produced by ten of the children in the *ESV* study. Finally, we address some of the theoretical issues which emerge from this preliminary analysis.

**2. Background**

Although all normally-developing children display comparable linguistic abilities at the earliest stages of life, language-specific influences soon become established in both perception and production. A great deal of research into children's speech has focused on the effects of different ambient languages on the development of phonological knowledge and phonetic performance (reviewed e.g. by Kuhl 1994, Vihman 1996).

Far less time has been devoted to investigating the effects on acquisition of structured variation within a single language – that is, the sort of variable features that have become a staple of much sociolinguistic work (see e.g. Chambers 1995). It is nonetheless widely assumed that children acquire their native accent at the same time as they acquire the abstract structures of their phonology. Chambers (1995: 158) summarises such a position as follows:

> 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 different 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.

While there is no denying that Chambers is right in this regard, two points must be considered. First, although children may speak with a recognisable accent, it is equally clear that there are differences between children's and adults' performance. Children do not produce fully mature speech patterns. 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. It is these two issues which drive our interest in the questions we raised in section 1 above. In short, what features comprise a child's accent, and at what rate are they acquired?

Early work in this area suggested that although a vernacular template is acquired in childhood, sociolinguistically-correlated variation emerges only in adolescence (Labov 1964). However, work by sociolinguists has led to a refutation of that view. Reid (1978), for instance, showed that pre-adolescents in Edinburgh are capable of displaying the same sort of style-shifting as are adults, and Romaine (1984) found gender-based correlation in phonological variables in children as young as 6. Similarly, ongoing work in Milton Keynes has highlighted differences between children at ages 4, 8 and 12 as accent features develop. Younger children tend to manifest patterns of speech behaviour closer to those of their parents, while older children become more subject to the influence of their peers (see e.g. Kerswill 1996, Kerswill & Williams 1997, Williams & Kerswill 1999).

Two studies have addressed the local speech patterns of children from the Tyneside area (the conurbation of which Newcastle upon Tyne forms the hub). First, Local (1978, 1983) studied a group of children aged between 4 and 6, and found that variation in vowel and prosodic patterns is both structured and changing as phonological rules and representations are mastered. More recently, Hartley's (1992) research with 5 year olds revealed gender-based variation in the realisation of glottalised forms of voiceless stops (see further Milroy, Milroy, Hartley & Walshaw 1994).

Only a very small number of sociolinguistic studies have 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 & Labov (1995) and Roberts (1997a, b) have shown that 3-year-olds in Philadelphia have already acquired complex patterns of vowel and consonant realisation, some of which are undergoing change in the local adult community. Roberts (1997a) furthermore found statistically significant differences between boys' and girls' usage of particular variants. Roberts & 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'.

Instrumental phonetic studies of children's speech have tended to look for evidence of particular phonological contrasts having been acquired, such as the voicing of English word-initial stops (see section 4.1). However, cross-linguistic investigations have revealed subtle differences in performance across different speaker groups from an early age. 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 & Buder 1994). Similar differences were found for vowel duration among the same children (Stoel-Gammon, Buder & Kehoe 1995).

It has also been shown that children learning Scottish English manifest dialect-specific features from at least the age of 6. Most varieties of English manifest vowel length differences conditioned by the voicing of a following consonant (e.g. Chen 1970). Dialects in Scotland, however, have a more complex distribution of long and short allophones, a patterning usually referred to as the Scottish Vowel Length Rule (SVLR; see Scobbie, Hewlett & Turk 1999). Analysis of Edinburgh children's speech has shown SVLR to be present by the age of 6, although the incidence is lower in cases where the children have one or more non-Scottish parent (Hewlett, Matthews & Scobbie 1999).[3]

[3] SVLR also occurs in Newcastle English (Milroy 1995), and will be investigated as part of the *ESV* project. A pilot study of a subset of the *ESV* children suggests SVLR is indeed present (Gaskell 1999).

To summarise this review of previous work, we can see evidence of dialect-specific features appearing from an early stage, and also evidence of changing patterns over the course of the acquisition period. One study suggests gender-based differences may appear by the age of 3. These are all issues of importance for the *ESV* project, to which we now turn.

## 3. The *ESV* project: methodology

### 3.1 Sampling

For the *ESV* project as a whole, two studies are being carried out in parallel. First, 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). The 8 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 neighbourhood 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 it is the mother who plays the larger role. The children are all monolinguals, and have no known speech or hearing disorders.

### 3.2 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*.

### 3.3 *ESV* variables

Recall that our aim in this project is to look at how 'sociophonetic' features emerge along with other systematic features of speech production. Our approach is therefore different from that taken in the majority of child-centred work in that we are not principally concerned with the emergence of phonological contrasts. Instead we are seeking to understand how children master the full range of alternants produced by adults. Analysis in *ESV* therefore centres on a range of the consonant and vowel variables which were investigated in the adult community as part of the *PVC* study. These include realisations of (t), as discussed in the present paper; linking and intrusive (r); glottalisation of (p) and (k); vowel length; 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 parents. Another future aim of the *ESV* project is to compare the phonological/phonetic features of child-directed speech with those of adult-to-adult speech.

## 4. Variable (t)

The research reported in the present paper concentrates on a subset of ten children from the cross-sectional study. Analysis of the full cohort is continuing, and will be reported in future publications.

The variable (t) is particularly complex in adult speech (see especially Docherty *et al.* 1997, and Docherty & Foulkes 1999a). First, adults use markedly different phonetic 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 effects are found, such that certain variants only occur in a restricted set of words. Specific details of these variants are given in sections 4.1 – 4.4 below.

The variable (t) has been analysed in four contexts, as shown in Table 1:

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

| 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 #* |

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 allophones of (t) produced by the children.

The following four sections (4.1 – 4.4) focus on these four 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 ten children's data. The entire tape recording for each child has been analysed for this purpose. Section 4.5 offers a summary of the findings.

### 4.1 Word-initial (t)

#### 4.1.1 Previous findings

Most varieties of English have voiceless stops that are aspirated in word-initial position. That is, the release of the oral closure is followed by a long lag in voice onset time (VOT). The minimum voicing lag required to signal a voiceless aspirated stop is 25 ms (Laver 1994: 349). 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 between 65 and 70 ms (e.g. Lisker & Abramson 1964; Kewley-Port & Preston 1974).

VOT of initial stops 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 & 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 & Preston 1974, Gilbert 1977, Smith 1978).

#### 4.1.2 *ESV* findings

All instances of word-initial /tV/ were located in the recordings. Following the established methodology, VOT duration was measured from the release burst of the stop to the onset of periodic noise signalling vocal fold vibration.

Table 2 presents the results of this study, showing mean durations, standard deviations and overall VOT ranges for each child. Pooled data for the ten children are also shown.

Table 2. VOT mean, s.d. and range for each child

| **child** | **sex** | **age** | ***N*** | **mean (*ms*)** | **st.dev.** | **range (*ms*)** |
|---|---|---|---|---|---|---|
| Naomi | f | 2;0 | *10* | 63 | 22 | 28 - 90 |
| Zack | m | 2;0 | *10* | 134 | 30 | 87 - 176 |
| Rayanne | f | 3;0 | *26* | 113 | 52 | 45 - 291 |
| Ryan | m | 3;0 | *39* | 84 | 32 | 27 - 175 |
| Owen | m | 3;0 | *22* | 95 | 34 | 35 - 144 |
| Liam | m | 3;0 | *29* | 64 | 26 | 21 - 144 |
| Rachel | f | 3;6 | *29* | 89 | 51 | 17 - 231 |
| Stephen | m | 3;6 | *21* | 102 | 55 | 35 - 276 |
| Hannah | f | 4;0 | *22* | 108 | 63 | 47 - 316 |
| Aimée | f | 4;0 | *7* | 86 | 55 | 32 - 186 |
| ***all*** | | | ***215*** | ***92*** | ***57*** | ***17 - 316*** |

The results of this study show close similarities with the previous studies of English-speaking children noted in section 4.1.1. The children are on the whole very variable in terms of the VOT length they produce, as demonstrated by the large standard deviations and the wide ranges displayed in Table 2. The average scores are also on the whole somewhat longer than those found for adults by Lisker & Abramson (1964), which is in keeping with general observations that segmental durations are longer in children's speech (Smith 1978).

However, almost all the tokens fall into the long lag VOT category (taking 25 ms to be the minimum appropriate value). Only 3 of the 215 tokens were shorter than 25 ms. Furthermore, 69% of the tokens fall into the 30 – 105 ms range which has been cited as typical for adults. These two findings suggest that the children have largely mastered the phonologically-governed variant appropriate for WI position.

### 4.2 Word-medial intersonorant (t)

#### 4.2.1 Previous findings

Auditory analysis of the *PVC* adults revealed that the majority of tokens in the WMIS context (e.g. *water*) in Newcastle English are glottalised. That is, they yield the auditory impression of glottal reinforcement of an oral stop. Data for the young WC adults, the group whose social characteristics are most similar to those of the *ESV* subjects, are shown in Table 3 (which represents a subset of the data in Table 9 of Docherty *et al.* 1997: 304). Glottalised forms, represented as [ʔ͡t],[4] were much the most common variant for both males and females, while few tokens failed to include any sort of glottal percept.

[4] Given the predominance of full voicing in such tokens, more revealing transcriptions might be used such as [d̬]. However, for the purposes of the present paper we employ the transcription which has been used most usually in discussions of Newcastle English.

Table 3. variants used in WMIS context, *PVC* young WC adults, expressed as %

| | t | ʔ͡t | ʔ | *N* |
|---|---|---|---|---|
| **females** | 10 | 77 | 13 | *245* |
| **males** | 4 | 88 | 8 | *191* |

Sociolinguistic differences also emerged among the *PVC* adults. Males in all groups – including the young WC – showed a significantly higher frequency of glottalised forms than the corresponding females.

Acoustic analysis of the glottalised tokens showed that 97% were produced with full or (less typically) partial voicing, and an interval of laryngealised voicing (Docherty & Foulkes 1999a, b). An example is shown in Figure 1, a spectrogram of the word *bottle* spoken by the mother of one of the *ESV* children.

### 4.2.2 *ESV* findings

It is not possible to draw direct comparisons between the *ESV* children and the *PVC* adults, since the two sets of data have been analysed in slightly different ways. However, the *PVC* data showed (i) that the great majority of WMIS tokens involved some sort of glottal articulation, and (ii) that almost all glottalised tokens are fully voiced with signs of laryngealisation. We therefore make a working assumption that the appropriate variant for children to target in this context involves both full voicing and laryngealisation.

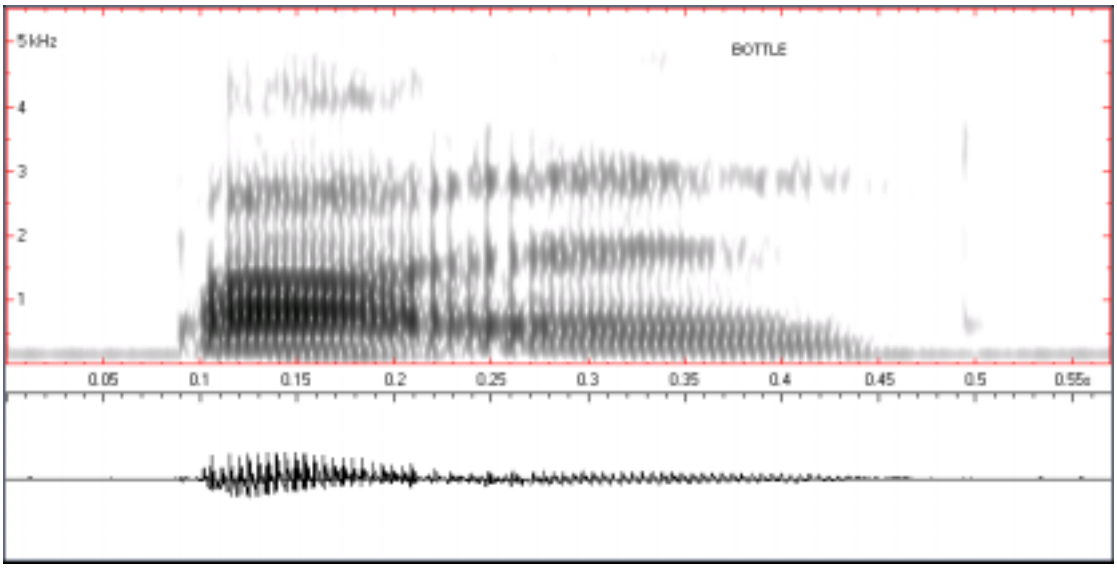
In order to examine the extent to which the salient features of the adult community are emerging in the speech of the children, we have therefore focused our acoustic analysis on the presence of full voicing and laryngealisation in the children's data. For these purposes, a token is defined as voiced if periodic noise is apparent throughout the target segment. Laryngealisation is defined as being present if it can be observed either during the target stop segment, or within the syllable nucleus either side of the stop (allowing for the fact that the adults also show variation in temporal co-ordination of the appropriate gestures; Docherty & Foulkes 1999a, b).

The findings of the acoustic analysis are shown in Table 4. Table 4 shows for each child the number of tokens which were fully voiced, and also those which displayed laryngealisation. (A third column indicates the number of 'pre-aspirated' tokens, an issue which is explained and discussed with reference to word-final pre-pausal tokens; see section 4.4.)

Table 4 suggests there is evidence that to some extent the children are reproducing key features found in the WC adult data, but to varying degrees. Laryngealisation occurs in under a third of cases in the children's speech, whereas it is overwhelmingly present in the adult tokens. There may also be signs of age-grading in the data shown in Table 4. The two youngest children have the lowest scores for laryngealisation, while in contrast, four year old Hannah produces 9 out of 15 tokens (60%) with laryngealisation. A spectrographic example from Hannah's recording is shown in Figure 2.

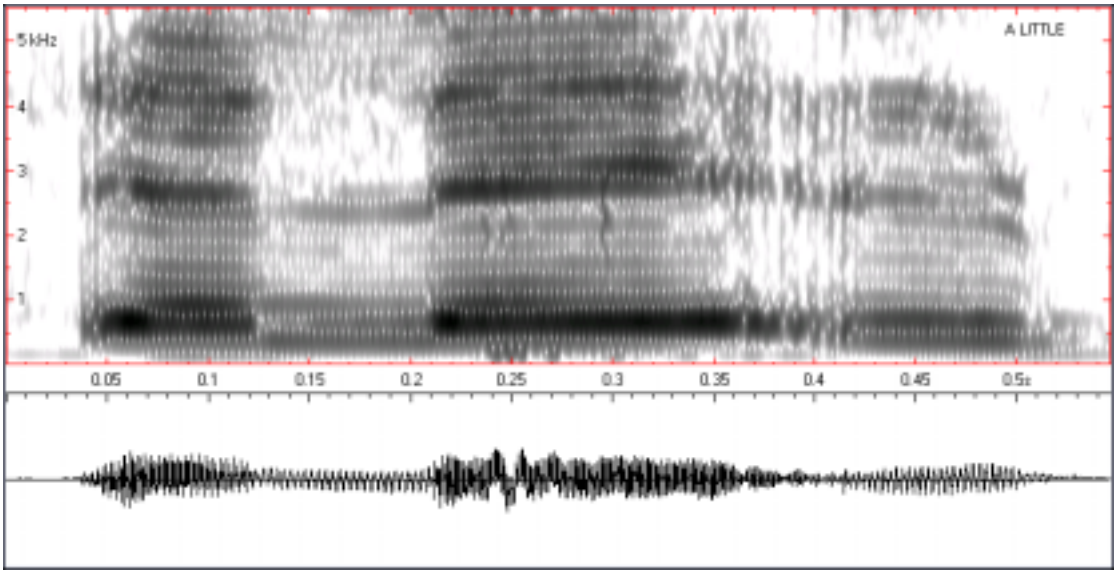
Fully voiced tokens account for around a quarter of the sample. While direct comparisons with adults are difficult to make, as already explained, this figure for the children appears relatively low, given the overwhelming presence in adult speech of glottalised tokens, which themselves tend to be cued by full voicing.

Figure 1. spectrogram of *bottle*, spoken by Stephen's mother. (t) is realised as a period of laryngealised voicing.




Click on the spectrogram to hear the sample

Figure 2. spectrogram of *a little*, spoken by Hannah. (t) is realised as a period of laryngealised voicing.




Click on the spectrogram to hear the sample

Table 4. acoustic patterns in WMIS context

| child | sex | age | *N* | laryngealised | voiced | pre-aspirated |
|---|---|---|---|---|---|---|
| Naomi | f | 2;0 | *16* | 2 | 7 | - |
| Zack | m | 2;0 | *11* | - | 5 | - |
| Rayanne | f | 3;0 | *14* | 10 | - | - |
| Ryan | m | 3;0 | *24* | 9 | 8 | 8 |
| Owen | m | 3;0 | *19* | 4 | 6 | 6 |
| Liam | m | 3;0 | *9* | 3 | - | - |
| Rachel | f | 3;6 | *38* | 14 | 9 | 2 |
| Stephen | m | 3;6 | *25* | 8 | 5 | 5 |
| Hannah | f | 4;0 | *15* | 9 | 8 | - |
| Aimée | f | 4;0 | *20* | 6 | 1 | - |
| ***all*** | | | ***191*** | ***57*** | ***49*** | ***21*** |
| ***all as %*** | | | | ***30*** | ***26*** | ***11*** |

### 4.3 Word-final intersonorant (t)

#### 4.3.1 Previous findings

As in word-medial position (section 4.2), word-final (t) may be glottalised or glottalled when it is immediately followed by a vowel. Thus, *get off* may be pronounced with the variant [ʔ͡t], or more rarely [ʔ]. In the *PVC* data, these variants accounted for 34% of the WFIS tokens produced by the young WC speakers, with no significant difference between males and females (Docherty *et al.* 1997: 293 [table 4]). There are, however, other variants which are possible in this context, including [ɾ], [t̠], [d], [ɹ], and [t]. The last variant, [t], accounted for only around 5% of the young WC sample, which enables us to generalise that 95% of tokens were voiced (assuming the glottal forms to be characterised by full voicing, as in section 4.2).

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 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]).

#### 4.3.2 *ESV* findings

In the *ESV* data the principal features sought in the acoustic analysis are therefore presence of full voicing and presence of laryngealisation. 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. (Again, 'pre-aspirated' tokens were also identified; see section 4.4 for discussion.) The acoustic patterns of the *ESV* analysis are summarised in Table 5.

Comparing Table 5 with Table 4, we find first of all that there are markedly fewer examples of (t) in WFIS context, reflecting the lower frequency of words in connected speech than in isolation. Two of the children (Naomi and Aimée) produced no examples at all of word-final (t) followed immediately by a vowel-initial word, while a third (Zack) produced only one example. The overall proportion of

laryngealised tokens, however, is very similar in the WMIS and WFIS contexts (30% versus 34% respectively). The data for the seven children who produce a substantial number of WFIS tokens are also fairly similar to their WMIS data in terms of the relative proportion of tokens which contain laryngealisation.

Table 5. acoustic patterns in WFIS context

| **child** | **sex** | **age** | ***N*** | **laryngealised** | **voiced** | **pre-aspirated** | **ɹ** |
|---|---|---|---|---|---|---|---|
| Naomi | f | 2;0 | - | - | - | - | - |
| Zack | m | 2;0 | *1* | - | - | - | - |
| Rayanne | f | 3;0 | *22* | 8 | 10 | - | - |
| Ryan | m | 3;0 | *14* | 4 | 7 | 1 | 1 |
| Owen | m | 3;0 | *23* | - | 22 | - | - |
| Liam | m | 3;0 | *6* | 3 | 1 | - | 1 |
| Rachel | f | 3;6 | *16* | 10 | 6 | - | 2 |
| Stephen | m | 3;6 | *5* | 2 | 1 | 1 | - |
| Hannah | f | 4;0 | *16* | 8 | 14 | - | - |
| Aimée | f | 4;0 | - | - | - | - | - |
| ***all*** | | | ***103*** | ***35*** | ***61*** | ***2*** | ***4*** |
| ***all as %*** | | | | ***34*** | ***59*** | ***2*** | ***4*** |

Clear differences emerge on comparison of the figures for voiced tokens. Although the immediate phonetic/phonological environment is constant, the children produce over twice as many voiced variants of (t) in word-final position (59% in Table 5) than in word-medial position (26% in Table 4).

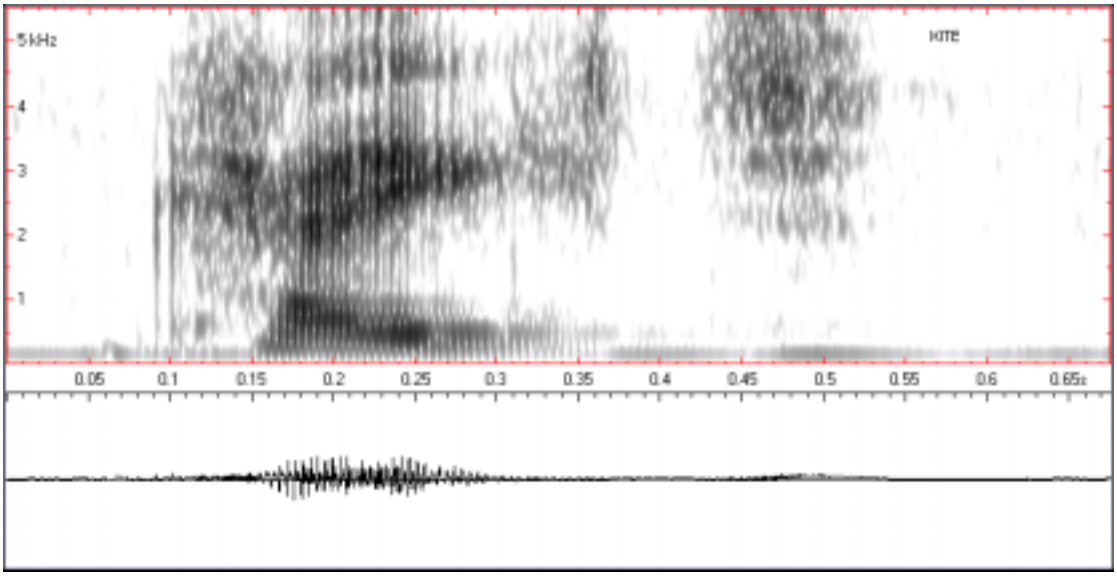
The [ɹ] variant is very rare in the children's data, occurring only four times in total, and in all cases being realised more like a labiodental or bilabial approximant, [ʋ]. However, all four cases do occur in appropriate lexical items (three examples in *got* and one in *get*), showing that the children who do use it are aware of the lexical selectivity of the variant.

### 4.4 Word-final pre-pausal (t)

#### 4.4.1 Previous findings

In contrast with the contexts described in 4.2 and 4.3 (as well as in contrast with most other British accents), pre-pausal (t) is almost never glottal(is)ed in Newcastle English. Common variants in this context include [$t^h$] and, for some speakers, pre-aspirated stops (Docherty & Foulkes 1999a). 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). An example from adult speech can be seen in Figure 3. This example of *kite* is produced by Hannah's mother, and shows a distinct portion of high frequency fricative energy prior to the stop closure.

Figure 3. spectrogram of *kite*, spoken by Hannah's mother. (t) shows pre-aspiration.




Click on the spectrogram to hear the sample

The rare glottal(is)ed forms in pre-pausal context invariably occur in a restricted set of words. Coincidentally, this set is very similar to that which accommodates the [ɹ] variant in WFIS context (section 4.3), i.e. common monosyllabic non-lexical words such as *what, it, that*, and common monosyllabic verbs such as *got* (Docherty *et al.* 1997). The acoustic correlates of these variants include creaky voice (as in section 4.2), which may be accompanied by patterns indicative of a glottal plosive articulation.

Both the pre-aspirated and glottal(is)ed 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 & 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 & Foulkes 1999a).

### 4.4.2 *ESV* findings

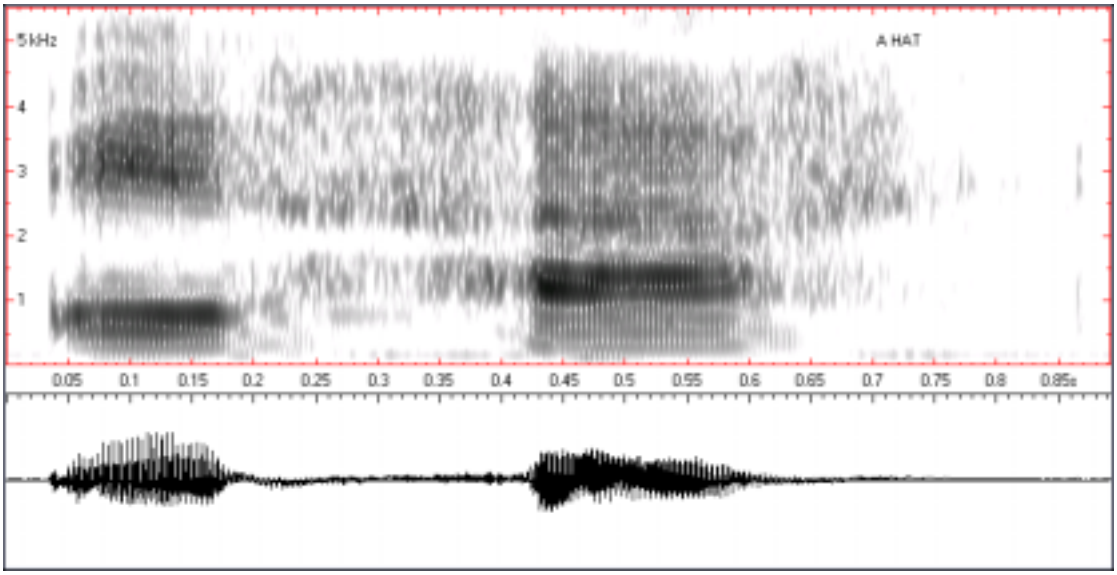
Analysis of the children's data has concentrated on seeking the presence of pre-aspiration and laryngealisation. A token is classified as pre-aspirated if it contains a period of distinct fricative energy prior to the stop closure, or a breathy voiced offset to the preceding vowel. To enable comparison with other contexts we also examined the presence of full voicing through the stop. Findings for WFPP position are shown in Table 6.

Table 6. acoustic patterns in WFPP context

| child | sex | age | *N* | laryngealised | voiced | pre-aspirated |
|---|---|---|---|---|---|---|
| Naomi | f | 2;0 | *10* | - | - | 4 |
| Zack | m | 2;0 | *8* | - | - | 8 |
| Rayanne | f | 3;0 | *51* | 21 | 1 | 15 |
| Ryan | m | 3;0 | *54* | 15 | 2 | 48 |
| Owen | m | 3;0 | *104* | 16 | - | 72 |
| Liam | m | 3;0 | *38* | 8 | - | 24 |
| Rachel | f | 3;6 | *55* | 4 | - | 37 |
| Stephen | m | 3;6 | *38* | 2 | 2 | 19 |
| Hannah | f | 4;0 | *24* | 4 | - | 14 |
| Aimée | f | 4;0 | *8* | - | - | 6 |
| ***all*** | | | ***390*** | ***70*** | ***5*** | ***247*** |
| ***all as %*** | | | | ***18*** | ***1*** | ***63*** |

As Table 6 shows, no less than 63% of tokens display pre-aspiration in pre-pausal position. Very few pre-aspirated tokens occurred in other contexts, as shown in Tables 4 and 5 above. Recall also that, in the adult community, this pattern is largely restricted to the speech of young WC women. The boys and the girls appear to be adopting this pattern in roughly equal measure. An example spoken by Hannah is given as Figure 4.

Figure 4. spectrogram of *a hat*, spoken by Hannah. (t) shows pre-aspiration.

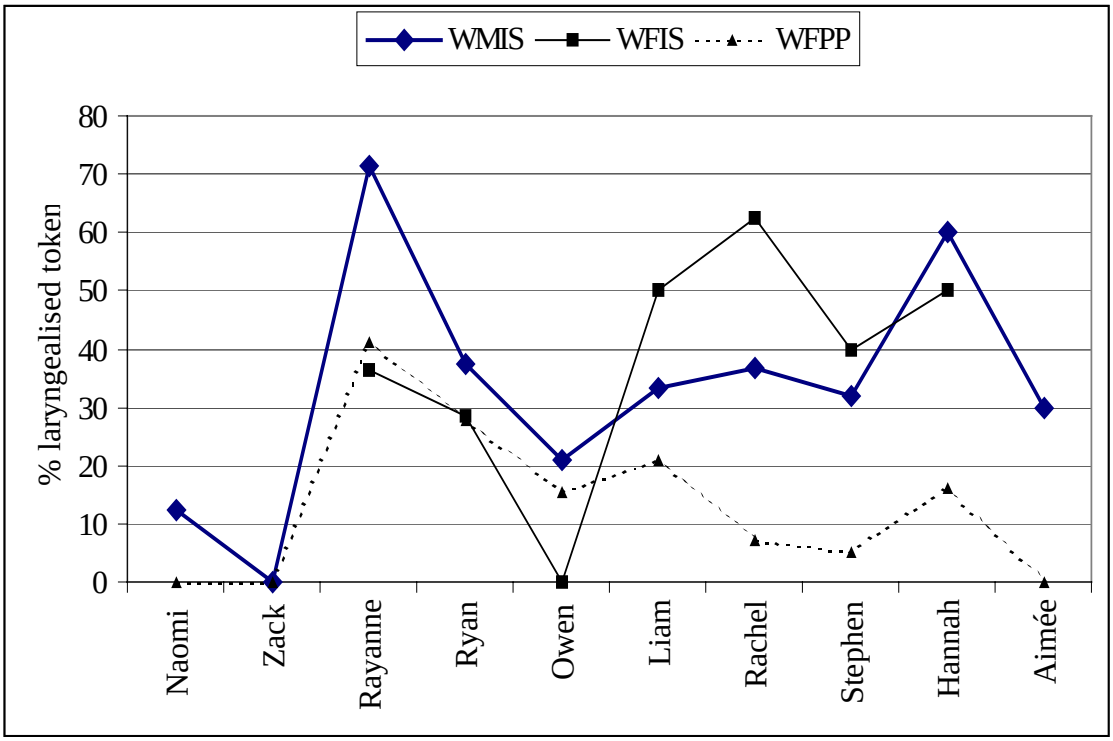



Click on the spectrogram to hear the sample

Voicing is overwhelmingly absent, as we would predict. Recalling that 59% of word-final tokens were fully voiced when followed by a vowel (Table 5), we can conclude that the children are in general sensitive to this phonologically-governed process of the dialect. Laryngealised tokens account for 18% of the total. Recall that in WMIS and WFIS contexts around one third of tokens involved laryngealisation. The bulk of the tokens are produced by just three of the children (Rayanne, Ryan, and Owen), and most children produce fewer laryngealised tokens in this context compared with WMIS and WFIS, as is appropriate following the adult model.

Figure 5 represents the proportion of laryngealised tokens produced by each child in the three contexts: WMIS (thick line), WFIS (thin line) and WFPP (broken line). The children are ordered by increasing age from left to right. For a child to have successfully mastered the adult patterns, we would expect high scores for WMIS, potentially high also for WFIS, but lower scores for WFPP. Figure 5 suggests that most of the children, particularly the older ones, are approaching an appropriate distribution of tokens, with the WFPP scores generally lower than those for the other two contexts. It is also noteworthy that 44 of these 70 tokens (63%) occur on just five lexical items: *what, got, that, not* and *it*. Recall that these are precisely the items in which glottal forms may appear, particularly in the speech of young WC women.

Figure 5. proportion of laryngealised tokens per child in three contexts.




Finally, our data as a whole show that the children's word-final variants match closely with those produced by adults: all children, even the two 2 year olds, show high degrees of pre-aspiration; in WFPP position the incidence of glottalisation is lower than in WFIS position for the oldest children; and the majority of glottal stop tokens occur in the same words which permit them in adult speech. These findings suggest that the children must be modelling their performance to a large extent on the adult forms. Note that this conclusion runs counter to frequent claims made on the

basis of similar observations of glottalisation (and deletion) in coda position in child speech, to the effect that children are operating a phonological rule of simplification (see e.g. Locke 1983: 230 ff.).

### 4.5 Summary of patterns

To recapitulate the patterns discussed in sections 4.1 – 4.4, we find:

- stops with long lag VOT in initial position
- a high proportion of voiceless pre-aspirated stops in pre-pausal position
- a range of variants in WMIS and WFIS positions which are not found initially or pre-pausally, including the salient local variants found in the adult community.

The children therefore appear to have made good progress in acquiring different allophones for the different phonological contexts. However, the match to adult forms is in some cases less successful than in others, notably in the WMIS and WFIS contexts. Laryngealisation is less common in the children's WMIS data than the adults', and similarly there are very few [ɹ]-like variants in WFIS compared with the *PVC* adults from the same community.

Where sociolinguistic differences exist in the adult community, the children's performance is again variable. Pre-aspiration is predominant in young women's speech, and also in the children's. By contrast, the use of the WFIS [ɹ] variant, which is also associated with young women, is largely absent from the children's data. 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 particularly marked differences are apparent in this sample of ten children.

Finally, the children show signs of sensitivity to the lexically restricted variants, [ɹ] in WFIS position, and [ʔ] in WFPP position. Although its overall rate of occurrence is low, [ɹ] always occurs in the words which tolerate it in adult speech. Similarly, around two thirds of the pre-pausal [ʔ] tokens occur in the few lexical items in which it may be found in adult speech.

## 5. Discussion

### 5.1 Developmental differences

In spite of the variability typical of children's speech, we have seen that our subjects in general produce appropriate phonetic forms in different phonological contexts. In some cases there is a high degree of consistency (e.g. the use of pre-aspirated forms in WFPP position and stops with long lag VOT in WI position). These findings lend support to those few previous studies which have been carried out in a similar vein, such as Roberts (1997a, b) and Hewlett, Matthews & Scobbie (1999).

The (t) variants we have concentrated on here involve subtle differences in the co-ordination of oral and laryngeal gestures, and it is the articulatory complexity involved in them which may go a long way towards explaining the differences between the children's performance and the adult model.

In initial position, a long lag VOT stop 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, Newcastle English demands a laryngealised stop. This comprises an

oral gesture similar to that for WI position, but with voicing throughout, and with a period of creaky voice timed usually to mask the oral release. With respect to the articulatory activity involved, the WMIS variant therefore differs in two ways from the WI variant: the vocal folds are adducted throughout rather than open at the beginning of the articulation, and a subsequent adjustment must be made to the laryngeal configuration in order to generate the creaky phonation (see e.g. Laver 1994: 330). In WFIS position this same variant may occur, as may various other voiced phones. The group of voiced phones includes [ɹ], which obviously 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. For a pre-aspirated stop, the voicing must be switched off *before* the oral closure is made.

The children's productions of (t) variants differ from those of the *PVC* adults in two main ways: the general absence of WFIS [ɹ], and the relatively low number of laryngealised tokens in WMIS position. The lack of [ɹ] may be explained by reference to both its physical and its phonological complexity. It is well documented that the approximant [ɹ] tends to be acquired late, because of the complexity of co-ordinating tongue retroflexion or bunching with both voicing and lip-rounding, and possibly also a pharyngeal constriction (Laver 1994: 302; Ladefoged & Maddieson 1996: 234). In the case of the [ɹ] variant used for (t) in Newcastle, there are also tight restrictions on the set of lexical items which may accommodate it in adult speech.

The requirement to adjust the laryngeal setting during the production of WMIS variants, and to time that adjustment with respect to the oral articulation, also presents a very complicated sequence for the children to learn. It is noticeable not only that the overall number of laryngealised tokens is low, but also that where they do occur there is often a fairly high degree of variation in the timing of laryngeal and supralaryngeal gestures compared with the *PVC* adults' productions. In WFIS position, then, it is no surprise that children tend to select articulatorily less complex variants from the range available. The majority of tokens used by the children involve full voicing, as is appropriate, but the children are less successful in grafting on the other gestures required for variants such as [ɹ] or [ʔ͡t].

There are also indications in our data that the children may be biased towards those patterns which are most in evidence in the speech of young women. At this stage of analysis we remain somewhat tentative in this conclusion, but the *ESV* data do appear to mirror the findings of Labov (1990) and Roberts (1997a, b). Labov (1990) accounts for such patterns with reference to the fact that, in general, children at this stage of development spend most of their time with women, in particular their mothers. It follows that they also receive the greatest part of their linguistic input from women, and are therefore likely to display speech patterns characteristic of a female model (but see Vihman 1993 for data which do not support this view).

### 5.2 The functionality of structured variation

Variability is one of the most obvious characteristics of children's speech. However, as a result of overriding methodological and theoretical concerns, linguists researching children's speech have tended to try and reduce this variability in order to ascertain whether children have successfully acquired more abstract (and presumably invariant) structures of the ambient language. Data such as those presented here demonstrate clearly that children are not impervious to structured variation in the ambient language. There is evidence that sociophonetic and allophonic aspects of

speech are learned alongside aspects usually considered as reflexes of the contrastive phonological system. There appears to be no difference between these processes, at least at first; both involve control of some of the same phonetic parameters at an equally fine-grained level. Variation in the phonetic substance that the children are exposed to is not simply discarded in the acquisition process – even at an early stage of development children are reproducing many of the salient features of the adult community.

This conclusion has implications for our understanding of how children build up a store of lexical representations, and of what form those representations take. In early language development, it is generally hypothesised that children operate with detailed holistic lexical representations (e.g. Studdert-Kennedy 1983, Pisoni 1993). The multiple trace model (Hintzman 1986, Jusczyk 1997) further allows different representations of single items to be stored. Phonemic or segmental awareness develops later, and may be linked to the onset of reading (Walley 1993). At this stage it is thought that the child may reorganise lexical representations, stripping away much of the detail of the early stages in favour of a more abstract representation based on phonological components. In other words, the multiple instances of a particular category converge on a small set of relevant categories (Jusczyk 1992, 1993). It is usually assumed that the details which characterised the early representation system become redundant, or are lost altogether.

If we follow models of this type, we can predict that the details encoded at the holistic stage of representation will include quite sophisticated allophonic material, and also features which are sociolinguistically relevant insofar as the adult community is concerned. There may 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 reorganised to reflect any sort of abstract phonological structure.

In light of this, we are persuaded to reconsider the role that structured variation might play in the acquisition process. Variability in the speech signal has typically been characterised by linguists as dysfunctional (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 characterised as dysfunctional with respect to acquisition. It is usually assumed that the main goal of the acquisition process is the learning of abstract contrastive units. These units must be extracted from their encoding within the phonetic medium. It follows, then, that the process of extraction ought to be more difficult when a unit is encoded in many different ways, e.g. in the form of different allophones. It is therefore commonly hypothesised that exposure to varying forms may delay or hinder the development of adult-like phonological categories (see e.g. Locke 1983: 202-3).

However, recent work in adult speech perception has identified a more positive role for variation. Although on-line lexical access may indeed be slowed down by variability in the signal, it has been shown that the task of learning phonological categories in a second language is improved by exposure to multiple voices (Lively, Logan & Pisoni 1993). Furthermore, exposure to a particular speaker's voice aids subsequent processing of speech produced by the same speaker. This implies that specific details may be recalled in the act of listening (Pisoni 1997), in turn suggesting that adults' cognitive representations of words may include highly detailed speaker-specific information alongside categorial features.

Our findings persuade us to take a similar stance on variation in first language acquisition. Our working hypothesis in the *ESV* project is to take the view that the child's task is not simply one of acquiring abstract phonological structures which underlie a superficially chaotic array of variant phonetic forms. Although the task of identifying these abstract structures is obviously a major part of the acquisition process, the learning of variable phonetic forms is also important as the child builds its sociolinguistic competence, to enable the child to use language appropriately in particular settings. These two tasks may appear in part incompatible. On the one hand, as we have discussed earlier, it is widely assumed that the extraction of abstract units may be rendered more difficult if the child is faced with variability in the speech signal; on the other hand, experience of structured variability is essential for the child to master allophonic, sociolinguistic and stylistic variation. Our data suggest that the *ESV* children have made good progress in replicating many aspects of the structured variability in adult speech, and they have achieved this before phonological competence is complete. This leads us to hypothesise that the process of gaining knowledge of structured variation may in fact help the child in the task of identifying abstract phonological structures, in much the same way that second language learners utilise variability to help them locate categories (Lively *et al*. 1993).

Specifically, recognising varying but recurrent phonetic patterns among exemplars of the same lexical items may serve to highlight the location of permutable components of words – e.g. phonemes, features, or gestures. As a consequence, the child may be helped in making the transition from the holistic stage of representation to one in which more abstract subcomponents are also identified.

How this process might work is represented schematically in Figure 6.

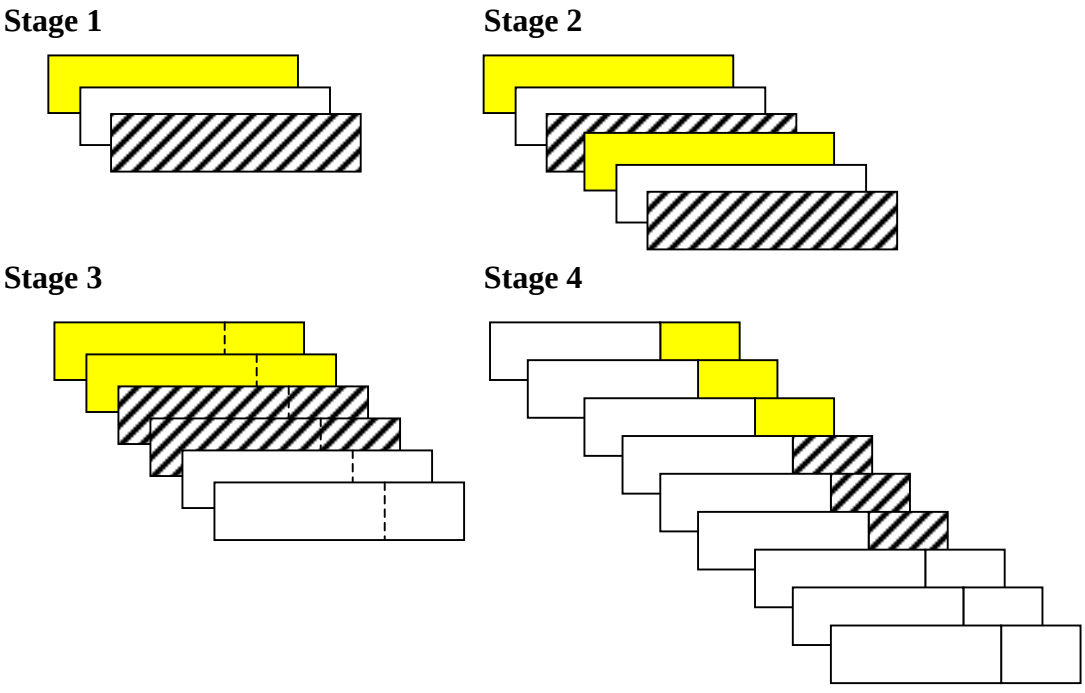



Figure 6. schematic representation of stages in the transition from holistic lexical representation to a system incorporating componential analysis.

At stage 1, a child may have heard a few exemplars of a particular word and stored them all holistically, unanalysed for componential structure. With further

experience, more exemplars are heard and memorised (stage 2), but patterns begin to emerge among the set of stored exemplars. These patterns may reflect, for example, versions produced by the same speaker, and/or, where structured variation occurs in the input, versions which contain a particular phone. By way of illustration, a child learning Newcastle English may hear versions of a word such as *get* which terminate in a variety of phonetic forms depending on context and/or speaker, including [tʰ], [ʰt], [ʔ͡t] and [ɹ].

At stage 3 analysis of the exemplars is beginning to take place. The store is becoming structured with reference to the repeated patterns. Groupings may be based, for example, on a particular speaker's voice, and/or they may reflect the presence of a particular phone within the exemplars. If particular phones are identified, we can see that the child is beginning to home in on the componential structure of the word. In the example of *get*, the final portion of the exemplars may contain any of [tʰ], [ʰt], [ʔ͡t] or [ɹ]. Identifying that the forms are lexically equivalent but variable within certain parameters enables the child to locate a boundary within the representation, in this example dividing the vocalic nucleus from the coda. This boundary is suggested by a dashed line in the stage 3 exemplars. A similar process involving different variants will lead eventually to a similar boundary placement between the other components of the word. Sufficient experience will then allow the child to ascertain the phonological, stylistic, and/or sociolinguistic roles of the particular variants. This in turn permits the child to use the variants appropriately, and will eventually see, for example, boys and girls diverging from each other in their speech patterns.

By stage 4 a more adult-like representation system emerges, with the child able to structure the store of tokens solely with reference to the sublexical components of the word. A more abstract superordinate representation may at this stage be erected to link all exemplars together – that is, a representation based on phonological units. The child has at this stage made the transition from holistic representation to a more abstract system. However, we suspend judgement as to whether the details characteristic of the earlier representations are lost at this stage. Some or all of the detailed information may remain accessible. Pisoni (1997) has discussed potential benefits for adults of lexical representations which are detailed rather than minimalistic. It strikes us that children would benefit in similar ways: they would be able to achieve fast recognition of speech from those speakers they have already heard, and they would furthermore be able to utilise their detailed knowledge to imitate and reproduce patterns they hear.

In this sketch of phonological acquisition phonetic details are not seen as an unwelcome complication, or excess baggage for lexical memory, to be discarded as soon as abstract forms have been extracted. Rather, recognition of patterns displaying structured variation may help in the identification of abstract units. The details themselves may then continue to hold a valuable status, and form an integral part of lexical representations even after analysis of phonological units has been achieved.

We might refer to the code-breaking process of locating components within stored representations as *triangulation*, by analogy with the surveying procedure of locating points in space with increasing accuracy as a function of the number of observations made from various directions. Comparison of different traces of a particular word from different 'angles' – i.e. involving different speakers and/or equivalent allophones – allows the child to identify the ranges within which smaller-scale units in the adult utterance may vary, thereby sharply resolving the internal phonological structure of word- or syllable-sized utterances.

## 6. Conclusion

The data presented here are preliminary, and our conclusions necessarily tentative. Ongoing work focuses on (t) productions of the remaining 30 children, and a longitudinal study of the speech of the youngest 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 the allophonic and socially-marked features which make up a child's developing sociolinguistic identity. Context- and accent-specific allophones are present very early, although some of the more physically complex (such as [ɹ], and the correct co-ordination of creaky phonation with oral gestures) seem to develop more slowly than others.

There may be evidence to support Labov's (1990) predictions of bias towards female paterns in the input, which has implications for the transmission of sound changes. The apparent closeness of the children's patterns to those of the 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 take as a point of departure the assumption that there are *language*-specific goals, rather than accent-specific ones, 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, J.K. (1995). *Sociolinguistic Theory*. Oxford: Blackwell.

Chen, M. (1970). Vowel length variation as a function of the voicing of consonant environment. *Phonetica* **22**. 129-159.

Docherty, G.J. & Foulkes, P. (1999a). Derby and Newcastle: instrumental phonetics and variationist studies. In Foulkes, P. & Docherty, G.J. (eds.) *Urban Voices*. London: Arnold. 47-71.

Docherty, G.J. & Foulkes, P. (1999b). Sociophonetic variation in 'glottals' in Newcastle English. *Proceedings of the XIVth International Congress of Phonetic Sciences*. University of California, Berkeley. 1037-1040.

Docherty, G.J., Foulkes, P., Milroy, J., Milroy, L. & Walshaw, D. (1997). Descriptive adequacy in phonology: a variationist perspective. *Journal of Linguistics* **33**. 275-310.

Gaskell, S.P. (1999). *An investigation into the presence and acquisition of Scottish Vowel Length Rule in the speech of three Tyneside children*. Unpublished undergraduate dissertation, University of Leeds.

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, S. (1992). *A study of the effect of sex and age on glottalisation patterns in the speech of Tyneside schoolchildren*. Unpublished undergraduate dissertation, University of Newcastle upon Tyne.

Hewlett, N., Matthews, B. & Scobbie, J.M. (1999). Vowel duration in Scottish English speaking children. *Proceedings of the XIVth International Congress of Phonetic Sciences*. University of California, Berkeley. 2157-2160.

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

Jusczyk, P. (1992). Developing phonological categories from the speech signal. In Ferguson, C.A., Menn, L. & Stoel-Gammon, C. (eds.) *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press. 17-64.

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

Jusczyk, P. (1997). *The Discovery of Spoken Language*. Cambridge, MA: MIT Press.

Kerswill, P. (1996). Children, adolescents, and language change. *Language Variation and Change* **8**. 177-202.

Kerswill, P. & Williams, A. (1997). Creating a new town koine: children and language change in Milton Keynes. *Reading Working Papers in Linguistics* **3**. 205-258.

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

Kuhl, P. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology* **4**. 812-822.

Johnson, K. (1997). Speech perception without speaker normalization. In Johnson, K. & Mullennix, J.W. (eds.) *Talker Variability in Speech Processing*. San Diego: Academic Press. 145-165.

Labov, W. (1964). Stages in the acquisition of standard English. In Shuy, R. (ed.) *Social Dialects and Language Learning*. Champaign, IL.: National Council of Teachers of English. 77-103.

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

Ladefoged, P. & Maddieson, I. (1996). *The Sounds of the World's Languages*. Oxford: Blackwell.

Laver, J. (1994). *Principles of Phonetics*. Cambridge: CUP.

Lisker, L. & Abramson, A.S. (1964). A cross-language study of voicing in intial stops: acoustical measurements. *Word* **20**. 384-422.

Lively, S.E., Logan, J.S. & Pisoni, D.B. (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, J.K. (1978). *Studies towards a description of the development and functioning of children's awareness of linguistic variability*. Unpublished PhD thesis, University of Newcastle upon Tyne.

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

Locke, J.L. (1983). *Phonological Acquisition and Change*. New York: Academic Press.

Macken, M.A. & Barton, D. (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, P.F. (1980). The control of speech production. In Yeni-Komshian, G.H., Kavanagh, J.F. & Ferguson, C.A. (eds.) *Child Phonology. Volume 1: Production*. New York: Academic Press. 9-21.

MacNeilage, P.F. (1997). Acquisition of speech. In Hardcastle, W.J. & Laver, J. (eds.) *The Handbook of Phonetic Sciences*. Oxford: Blackwell. 301-332.

Milroy, J. (1995). Investigating the Scottish Vowel Length Rule in a Northumbrian dialect. *Newcastle and Durham Working Papers in Linguistics* **3**. 187-196.

Milroy, J., Milroy, L., Hartley, S. & Walshaw, D. (1994). Glottal stops and Tyneside glottalization: competing patterns of variation and change in British English. *Language Variation and Change* **6**. 327-357.

Milroy, L., Milroy, J., Docherty, G.J., Foulkes, P. & Walshaw, D. (1999). Phonological variation and change in contemporary English: evidence from Newcastle upon Tyne and Derby. In Condé Silvestre, J.C. & Hernández-Campoy, J.M. (eds.) *Variation and Linguistic Change in English*. Cuadernos de Filología Inglesa **8**. 35-46.

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

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

Reid, E. (1978). Social and stylistic variation in the speech of children: some evidence from Edinburgh. In Trudgill, P.J. (ed.) *Sociolinguistic Patterns in British English.* London: Arnold. 158-171

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

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

Roberts, J. & Labov, W. (1995). Learning to talk Philadelphian. *Language Variation and Change* **7**. 101-112.

Romaine, S. (1978). Postvocalic /r/ in Scottish English: sound change in progress? In Trudgill, P.J. (ed.) *Sociolinguistic Patterns in British English.* London: Arnold. 144-157.

Scobbie, J. M., Hewlett, N. & Turk, A. (1999). Standard English in Edinburgh and Glasgow: the Scottish vowel length rule revealed. In Foulkes, P. & Docherty, G.J. (eds.) *Urban Voices*. London: Arnold. 230-245.

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

Stoel-Gammon, C., Buder, E.H. & Kehoe, M.H. (1995). Acquisition of vowel duration: a comparison of Swedish and English. *Proceedings of the XIIIth International Congress of Phonetic Sciences* **4**. 30-36.

Stoel-Gammon, C., Williams, K. & Buder, E.H. (1994). Cross-language differences in phonological acquisition: Swedish and American /t/. *Phonetica* **51**. 146-158.

Studdert-Kennedy, M. (1983). On learning to speak. *Human Neurobiology* **2**. 191-195.

Vihman, M.M. (1993). Variable paths to early word production. *Journal of Phonetics* **21**. 61-82.

Vihman, M.M. (1996). *Phonological Development*. Oxford: Blackwell.
Walley, A.C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review* **13**. 286-350.
Watt, D.J.L. & Milroy, L. (1999). Patterns of variation and change in three Newcastle vowels: is this dialect levelling? In Foulkes, P. & Docherty, G.J. (eds.) *Urban Voices*. London: Arnold. 25-46.
Williams, A. & Kerswill, P. (1999). Dialect levelling: continuity vs. change in Milton Keynes, Reading and Hull. In Foulkes, P. & Docherty, G.J. (eds.) *Urban Voices*. London: Arnold. 141-162.
Wolfram, W. (1989). Structural variability in phonological development: final nasals in Vernacular Black English. In Fasold, R. & Schiffrin, D. (eds.) *Language Change and Variation*. Amsterdam: John Benjamins. 301-332.

*Paul Foulkes, Dominic Watt*
*Department of Linguistics and Phonetics*
*University of Leeds*
*Leeds*
*LS2 9JT*

*p.foulkes@leeds.ac.uk*
*d.j.l.watt@leeds.ac.uk*

*Gerry Docherty*
*Department of Speech*
*University of Newcastle*
*Newcastle upon Tyne*
*NE1 7RU*

*g.j.docherty@ncl.ac.uk*