Phonetic devices and the construction of the phonological space

Elinor Payne\textsuperscript{1}, Brechtje Post\textsuperscript{2}, Nina Gram Garmann\textsuperscript{3}, Hanne Gram Simonsen\textsuperscript{4}
\textsuperscript{1}Oxford University, \textsuperscript{2}Cambridge University, \textsuperscript{3}Oslo & Akershus University College, \textsuperscript{4}Oslo University
elinor.payne@phon.ox.ac.uk

When learning to speak, infants’ exposure to the distributional frequencies (in e.g. syllable structure, etc.) of the ambient language can subtly influence the developmental path followed (Vihman & Velleman, 2000; Prieto et al, 2006) despite common neurophysical constraints as the infant capitalizes on that language’s statistical properties to learn its structures (Saffran et al, 2003). Infants have to learn the language-specific structure, and also have to acquire and fine-tune the necessary phonetic skills to be able to map these structures onto speech output. But languages also differ in the way structure is phonetically implemented, making different use of a range of ‘phonetic devices’, and children also have to learn both the appropriate phonetic devices and how these are correctly mapped. A particular phonetic ‘device’ may also be a ‘multiple signifier’: e.g., the relative timing of VC sequences is known to be an important cue to coda obstruent voicing in English (e.g. beat/bead), but the duration of the vowel in itself may also be an associative cue to vowel quality (e.g. bit/beat). Similarly, in Norwegian, VC timing is known to play some role in the cueing of [voice] in coda obstruents, but is also a key marker of syllable weight (e.g. hat/hatt). Therefore, while there may be superficial similarities between languages in the phonological distinctions made, and in the phonetic devices employed to mark these, the ecology of available devices and the different demands placed on these will vary. This paper examines the degree to which such different ecologies impact the acquisition pathway, by comparing the acquisition of VC timing in Norwegian and English.

We recorded and analysed productions of 9 English-learning and 9 Norwegian-learning children aged 2;6, 4 and 6 years, elicited via a picture-naming task which prompted monosyllabic words containing the relevant features in both languages (e.g. back, bag, feet, head, sit, lid; bok, smokk, dag, flagg), and compared these with adult productions. We found that in both languages the youngest children already use different VC timing relations (as expressed by vocalic-ness of the rhyme, or V/VC) to distinguish fortis and lenis rhymes, though the distinction increases with age. In English, as with adults, the mean duration of V is always less than 50% of total VC in fortis rhymes, and always more than 50% in lenis rhymes, for a range of vowel and consonant types (see Figure 1a). In Norwegian, fortis/lenis timing interacts with rhyme length contrast to produce 4 categories of timing, and only short vowels are less than 50% of total VC in fortis rhymes, and only long vowels are more than 50% in lenis rhymes (see Figure 1b). This would appear to suggest that already at the youngest age the children have the relevant phonological distinctions in their production plan, and are mapping them onto phonetic devices (VC timing relations) in a broadly language-appropriate manner, even though absolute durations of vowels and consonants are overly long in both languages.

\textbf{Figures 1a and 1b: V/VC in fortis and lenis rhymes in English (left) and Norwegian (right)}

Scrutinising the child productions further, it emerges that the precise details as to how these timing relations are achieved vary. For example, in English, appropriate degree of vocalic-ness in the rhyme is achieved through complementary timing (longer vowels precede shorter lenis stops, shorter vowels precede longer fortis stops), and this production strategy is observable in child speech from the age of 4 years. At 2;6 years, however, the target vocalic-ness of the rhyme is achieved through modification to the duration of \textit{either} the consonant \textit{or} the vowel, seemingly as a function of vowel and consonant type (see Figures 2a and 2b). We attribute this behaviour to immaturity in phonetic co-ordination at this age, and the asymmetry of strategies adopted to the differing articulatory demands associated with
different vowel heights and consonant place of articulation.

**Figures 2a and 2b: mean absolute V and C durations in English for different VC combinations**

Thus, at 2;6 years, despite immature phonetic ability, mappings between phonological structures and phonetic devices are approximately adult-like (at least in relational terms). However, the demands of temporal co-ordination become more complex and challenging as speech becomes faster and more connected, resulting in mappings that undermine the contrasts being expressed. Figures 3a and 3b show mean vocalic-ness (V/VC) for both Norwegian V:C and VC: rhymes, for sonorants and voiceless stops (blue = 2;6 yrs; green = 4 yrs; beige = 6 yrs; purple = adult). It can be seen that V/VC is a stable differentiator of rhyme type, and close to the adult target, at all ages, except at 6 years, where there is no significant difference between the two rhyme types. We conclude that the more crowded ecology of demands for VC timing in Norwegian has a greater impact on Norwegian-learning children than on English-learning children at this age, resulting in an apparent U-shaped curve in development (temporary regression from earlier adult-like structures).

**Figures 3a and 3b: mean V/VC values for V:C and VC: rhymes in Norwegian**

In English the temporal relations of VC appear to be maintained more robustly over the developmental arc. There are, however, potential knock-on effects elsewhere, e.g. in the use of duration to mark tense/lax distinctions in vowels. We find early use of duration differences in pairs like beat/bit and bead/bid, but these are no longer robust and have been overtaken by spectral differences by age 4, indicating a shift in priorities between competing uses of durational cues.

These findings suggest the acquisition pathway is determined not just by developing articulatory skills and knowledge of a particular phonological system, but also by a play-off between competing uses for duration as a linguistic-phonetic ‘device’ for implementing that system. In learning to speak, even once children are aware of phonological distinctions and have the phonetic skills to implement them, they still have to negotiate the complex mapping between the two. Our study suggests these mappings themselves take some time to fall into place, and competition between mappings may be a contributing factor to the time and pathway taken.

We discuss the implications of these findings for the acquisition and dynamic nature of the phonological representations themselves. If the phonetic exponent, and the mapping between phonological structure and that exponent, are in a state of flux, to what extent must the representations themselves also be dynamic? We argue that phonological representations are not acquired in isolation, but emerge as part of a complete, dynamic system. Each target language to be acquired presents a unique set of structural relations and ecology of phonetic devices to signal these. During acquisition, by exploring the possible mappings and how they may be in competition, the child both seeks out the phonetic shape as projected by the phonological space of her ambient language, and constructs that phonological space for herself.