Background

The best substantiated of Jakobson’s (1941) claims about implicational universals in acquisition concerns phonetic type: Voiceless unaspirated are produced first, and the mastery of voicing contrasts in stops implies the mastery of voiceless unaspirated stops. Claims in the literature in studies of many languages:

1) unaspirated stops are produced in babbling at 6-7 months
2) unaspirated stops mastered about 2 years in English (MacKen & Bartos, 1985), Cantonese (Chamlee et al., 1983), etc.
3) voiceless stops are not mastered until 4-5 years in French (Allen, 1981), Thai (Gandour et al., 1986), Hindi (Davis, 1993), etc.

Validation depends on accurate phonetic characterization: e.g., if English stops in word-initial position characterized as in traditional accounts (i.e., voiced [b, d, g] in opposition to voiceless [p, t, k]), then English is exceptional in claim.

Therefore, it is important to take language-specific phonetic characteristics into account in comparing the acquisition of the “same” phonetic types across languages.

Questions for this study

What are the implications for Jakobson’s developmental universal of more recent phonetic studies documenting much finer-grained phonetic differences across languages having purportedly identical phonetic categories?

1) Is VOT the only phonetic way of describing the phonetic types of a given language?

Can we use a combination of VOT values and other measures in children’s productions and adults’ productions to explain the developmental universal?

Data

Word-initial coronal and dorsal stops in four languages with a two-way phonetic type contrast, elicited in a word repetition task using an audio prompt only (or for adults) or using an audio prompt accompanying a picture (for children).

Three words were used for each target in a variety of word contexts (aiming for five to seven contexts for each target). For example:


Subjects were about 20 children aged 2-5 years and 3-5 adult speakers for each language.

Analysis 1 — VOT

We measured voice onset time (VOT), the latency from the burst that marked the release of the lingual constriction to the onset of voicing in the closure (negative VOT — “voicing lead”) or in the following vowel (positive VOT — “voicing lag”).

Histograms of all of the adult values (Fig. 1, left panel) showed expected patterns of (1) short vs. long lag values for Cantonese and (2) mostly short vs. long lag values for English.

• However, (3) the voiceless stops of Japanese had “short” lag values that are intermediate between the truly short lag values for the unaspirated stops and the long lag values of the aspirated stops of Cantonese. Moreover, (4) the voiced stops of Japanese showed a bimodal distribution, with many short lag values that overlapped with the values for the voiceless stops. This was particularly true of the dorsal stops (Fig. 2, left panel).

• By contrast, (5) the voiced stops of Greek showed much larger voiceless (leading voice) values which did not overlap at all with the values for the voiceless stops, even though (6) the voiceless stops showed prototypical short lag values.

• The children’s VOT patterns (Figs. 1-2, right panels) showed distributions very much like the adult patterns, even for minimally language-specific details such as (1) the intermediate lag values of the Japanese voiced stops and (5) the particularly long voice-leading values of Greek.

(Q1) What is the acoustic cue that separates the Japanese voiced from voiceless stops in the region of overlapping VOT values?

Note that this pattern of overlapping values is superficially like the VOT distribution in English children with “ covert contrast” (MacKen & Barton, 1980).

(Q2) Why are the Greek children capable of producing the very long voice-leading characteristics of the Greek voiced stops?

Recall that these children are all much younger than the ages reported for mastery of voiceless stops in French, Thai, Hindi, etc.

Analysis 2 — Burst intensity

To answer Question 1, we took a measure of burst intensity that was the sum of energy in the region above 2000 Hz in a spectrum calculated over a 5 ms window beginning at the stop release. We thank the speakers who let us record them.

This work was supported by NIH grant 02932 to Jan Edwards.

Figure 3. Burst intensity as a function of VOT for Japanese voiced (red circles) and voiceless (yellow triangles) stops. Left panel shows the entire distribution, right panels zoom to the region of overlapping VOT values.

• MANOVA analyses showed a significant (p = 0.001) difference between the voiceless and voiced stops in the region of VOT overlap for both places of articulation for both the adult productions (Fig. 3a) and the child productions (Fig. 3b).

• Given the greater intra-vocal tract variation build-up in voiceless stops, we predicted greater burst intensity in the voiced stops. The mean difference in burst intensity was in the predicted direction.

Discussion and conclusion

The four languages that we examined in this study all have been described in terms of a simple two-way phonation-type contrast between either short lag and long lag (Cantonese and English) or between voiceless and voiceless (Greek and Japanese).

However, a closer examination of the acoustic phonetic contrasts in Greek and Japanese suggests that more is involved than control of VOT, and that the other parameters involved facilitate the mastery of this difficult or “marked” contrast further lent further support to an interpretation of Jakobson’s developmental universal in terms of the universal constraints imposed by the more or less stringent articulatory and aerodynamic requirements of the more or less “marked” sounds.

Acknowledgements

This work was supported by NIDCD grant 02932 Jan Edwards. We thank the speakers who let us record them.