AN ULTRASOUND INVESTIGATION OF POSSIBLE COVERT CONTRASTS IN FIRST LANGUAGE ACQUISITION: THE CASE OF /sp ~ sw ~ sm/ > [f] MERGERS

Sonya Bird
Dept. of Linguistics, University of Victoria, PO Box 3045, Victoria BC, V8T3K5 sbird@uvic.a

1. INTRODUCTION

Child language often involves consonant cluster simplification. In the case of clusters beginning with /s/, e.g. /sp/, the most common strategy is to delete /s/, resulting in forms like [pun] for *spoon* (Ingram 1991). An alternative strategy attested in some children is to merge the two consonants, e.g. /sp/ > [f], resulting in [fun] for *spoon*. (Locke 1983). Mergers provide interesting insights into phonological representation in children: they suggest that children are able to decompose sounds into their composite features. Indeed, in the example above, [f] is not a compromise between /s/ and /p/ along any single phonetic continuum; rather it is a merger of the manner of articulation of /s/ with the (near-) place of articulation of /p/.

To fully understand the implications of mergers for phonological representation in children, it is essential to consider the phonetic detail of these mergers. Specifically, is the result of merger truly identical to the underlying version of the sound? In the above example, is the perceived [f] in *spoon* identical to the /f/ in *food*, or does it retain some trace of the underlying segments /sp/? This question has been difficult to answer because the relevant data from child language is based primarily on auditory analyses (in the form of transcriptions), which reflect what the listener hears rather than what the speaker is doing. As a result it is not clear whether perceived mergers are truly complete from the speaker's perspective (Scobbie et al. 2000). As a step towards overcoming this problem, this paper reports on an preliminary articulatory study conducted with a single, normally developing, child whose pronunciation of /s/ + labial consonant sequences (e.g. /sp/) is not auditorily distinct from her pronunciation of /f/. Results suggest that in her case, the mergers are indeed complete.

2. METHOD

The experiment was conducted in two sessions, separated by approximately one month. The participant for the study was a single, normally developing girl – AC – who was 2;8 years at the first session and 2;9 years at the second (her pronunciations did not change across sessions). In both sessions AC was recorded using lingual ultrasound (GE Logic E portable ultrasound machine with 8C-RS convex transducer) pronouncing randomly ordered words from sets like *sponge* – *sun* – *fun*. Table 1 provides the stimuli list, which included only words familiar to AC (*-ed words were elicited in the second session only).

During the recordings, AC was seated on her mother's lap, with the ultrasound machine to the left. AC's mother held AC's head stable against her chest with one hand; she held the transducer under AC's chin with the other hand, monitoring the screen of the ultrasound machine to ensure the probe remained immobilized. AC's father sat to her right, and elicited the target words by asking her to name pictures he presented to her one at a time. The sessions were also audio-taped, the first session using a Countryman AKG C429 head mounted microphone resting on the stand of ultrasound machine (being too big to place on AC's head); the second session using a Senheiser ME-55 directional microphone (with a K6 capsule). Ultrasound and audio recordings were mixed through a Mackie 1402-VLZ3 mixer and captured on an external PC computer using Sony Vegas Pro 8. Because of technical difficulties in the first session, the recording was repeated 3 times, such that the first two repetitions acted as practice sessions. This did not appear to affect the pronunciation of the target words. In the first session, only the midsagittal view of the tongue was recorded; in the second session, the recording was done twice, to get a cross-sectional as well as a midsagittal view of the tongue (by turning the probe 90 degrees).

Table 1. Experimental stimuli.

V	/sw~sp~sm/	/s/	/f/
[I]	swimming; swing; swing set	sing; singing	finger
[u]	spoon; smoothy	soup *	food
[ɛ]	sweater *	sesame *	fender *
[□]	Sponge Bob	suns; sunglasses	fun

Figure 1 illustrates the experimental set-up. Although it did not provide the kind of controlled stabilization required for rigorous, quantitative analysis of the data, it was deemed sufficient to gain a qualitative understanding of tongue position and movement during the target sounds. Certainly a more complete follow-up study should include more reliable head and probe stabilization, to the extent that this is possible with such a young participant.

Fig. 1. Experimental set-up.

3. RESULTS

In total, 40 tokens were elicited in the two sessions, based on a midsagittal (from the side) view of the tongue (AC repeated some of the tokens; these were kept in the analysis): 12 of /s/, 8 of /f/, 11 of /sw/, 6 of /sp/ and 3 of /sm/. Using the audio signal as a reference, short ultrasound (video) clips corresponding to the target sounds were inspected visually to see how /sw~sp~sm/ compared to /s/ vs. /f/ in terms of tongue position. Because the tongue contour was not always clear, and because the probe angle was not consistent within and across sessions, judgments on tongue position were not always easy to make. However in general, results indicate that /sw~sp~sm/ are identical to /f/: they were consistently transcribed as [f], with the exception of two tokens of /sw/, transcribed as [f^w]. Articulatorily, the most reliable difference between /sp~sw~sm~f/ and /f/ vs. /s/ was tongue tip (TT) position: /sw~sp~sm/ and /f/ never showed clear TT raising; while tongue contours for /s/ were often unclear, they suggested TT raising in 9/12 tokens.

Figure 2 provides illustrations of this overall pattern. In each frame, the white line represents the tongue contour, with the tongue tip on the right. Panel 2a is of *swing set*; it shows the beginning and end frames of the /sw/ cluster followed by frames from /I/ and the initial /s/ of 'set'. These frames show that /sw/ is relatively stable throughout its duration (1st and 2nd frames); it does not exhibit a raised TT like the following /s/ (4th frame), nor does it exhibit tongue back raising as one might expect for /w/ (tongue position is stable from /sw/ to /I/ - 2nd and 3rd frames). Panel 2b is from *finger* and shows that /f/ is also stable throughout its articulation, and without TT raising. Panel 2c is from *a sun* and shows TT raising for the initial /s/ (2nd and 3rd frames).

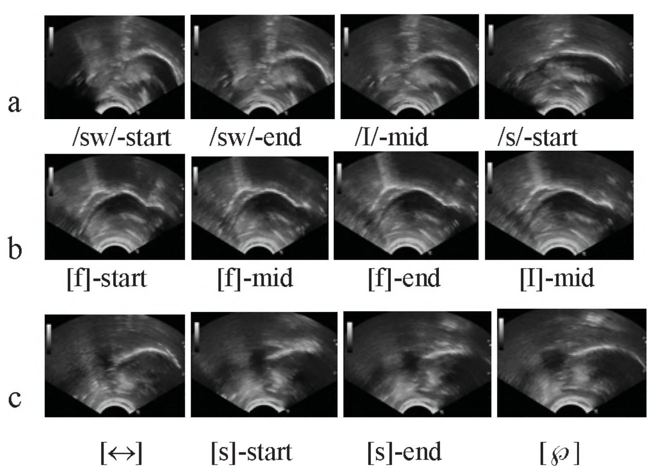

Fig. 2. Illustrative examples – midsagittal view. Panel a: *swingset;* panel b: *finger*; panel c: *a sun*.

During the second session target words were elicited using a cross-sectional (from the front) view of the tongue as well as a midsagittal view, to investigate tongue grooving. In total, 17 tokens were elicited this way: 5 of /s/, 5 of /f/, 3 of /sw/, 2 of /sp/ and 2 of /sm/. In addition, sustained, isolated /s/ and /f/ tokens were elicited. Unfortunately, none of the 5 /s/ tokens elicited in real words had clear tongue contours, making it impossible to compare initial consonants across sets of words. Again, results suggest that /sw~sp~sm/ are identical to /f/: aside from a single frame in a single /sp/ token that exhibited /s/-like grooving (see Fig. 3b, 1st frame), all /sw~sp~sm/ tokens looked like /f/: they had either no tongue grooving or a shallow wide grove, in contrast to the narrower, deeper groove of the sustained /s/. Figure 3 provides typical examples: Panel 3a is of sustained /s/ and two versions of /f/, with and without grooving; panel 3b is of /sp/ in *Sponge Bob* (/s/-like), sweater (/f/-like, with groove) and *spoon* (/f/-like, no groove).

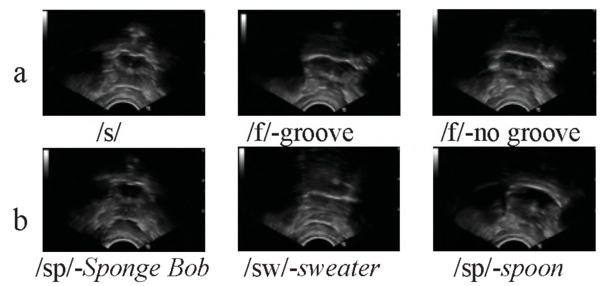

Fig. 2. Illustrative examples – cross-sectional view. Panel a: sustained /s/ vs. /f/; panel b: /sp/ in 3 different tokens.

4. DISCUSSION

If some trace remained of the two underlying segments in /sw~sp~sm/ clusters, one might argue that the observed merger is simply a phonetic effect, a result of AC's developing motor skills. The results tentatively indicate that this is not the case: the /sw~sp~sm/ > [f] merger is complete. For this reason, it is proposed that the merger is a phonological effect, because it can only occur if AC has access to some level of abstraction. Indeed, it is only in abstraction that [f] consists of a compromise between /s/ and /p/, incorporating the manner of articulation of /s/ (fricative) with the place of labial consonants. Crucially, [f] does not represent a compromise between /s/ and /w~p~m/ along any single phonetic continuum. In terms of place for instance, the primary articulators involved are different for /s/ vs. /w~p~m/: the tongue vs. the lips. Therefore, without the ability to abstractly decompose sounds into their place and manner features, there would be no reason to come up with [f] as a merged version of / sw~sp~sm/.

REFERENCES

Ingram, D. (1991) *First language acquisition: method, description and explanation*. Cambridge: Cambridge University Press.

Locke, J. (1983) *Phonological acquisition and change*. New York: Academic Press.

Scobbie J., F. Gibbon, W. Hardcastle & P. Fletcher (2000). Covert contrasts as a stage in the acquisition of phonetics and phonology. In M.B. Broe & J.B. Pierrehumbert (Eds) *Papers in Laboratory Phonology V: Acquisition and the lexicon*. Cambridge: Cambridge University Press.