Language-Specific and Language-Universal Aspects of Lingual Obstruent Productions in Japanese-Acquiring Children

日本語獲得における幼児の舌阻害音生成の普遍性と個別性

Mary E. BECKMAN*, Kiyoko YONEYAMA**, and Jan EDWARDS*

*オハイオ州立大学教授 (Professor, Ohio State University)

**大東文化大学講師 (Assistant Professor, Daito Bunka University)

Keywords: alveolar backing, covert contrast, developmental universals, markedness, constraints, phoneme frequency, phonological acquisition, stopping of fricatives

要旨：音韻獲得のメカニズムは言語普遍的であり、音素目録を形成する多くの言語普遍的な原理によって決定されるという Jakobson （1941/1968）の主張は、現在の音韻獲得理論の基盤となっている。しかしながら、近年の多言語にわたる音韻獲得に関する研究結果から、言語普遍的な原理だけでは全ての言語の音韻獲得過程が説明できるわけではないことも明らかになってきた（e.g., Ingram, 1999; Vihman, 1993; Vihman et al., 1985）。本稿では日本語を獲得する幼児の舌阻害音の生成過程についての研究結果を英語と比較対照することにより、日英語における舌阻害音の音韻獲得順序の相違が両言語の個々の音とその音の組み合わせの頻度に帰依している可能性が高いことを示した。また、調音の難易度といった言語普遍的側面と言語によって異なる個々の音とその音の組み合わせの頻度という言語個別側面の双方を考慮した音韻獲得理論の必要性について論じた。

1. Variability and Continuity in Phonological Acquisition

In his influential monograph on child language, aphasia, and phonological universals, Jakobson (1941/1968) made several claims that continue to shape how we think about acquisition today. The most general claim was that there are universal substantive principles that structure the phoneme inventories of all spoken languages, and that these principles also determine how children acquire speech. More specifically, Jakobson assumed that there is a clear discontinuity between the rich set of consonant- and vowel-like sounds that children produce randomly in their preverbal babbling and the initially impoverished but reliable inventory of phonemes that they produce in their first words. Moreover, the consonants and vowels that are first to be mastered are the same for all children everywhere, and each child expands this initial inventory in a rigid universal order.

Both of the more specific claims are wrong. Arguing against a constant order of acquisition is mounting evidence of variability. For example, Ingram (1999) summarizes studies of five or more children acquiring each of English, Quiché, Turkish, and Dutch to show that "typical" consonant inventories at 20-27 months differ for these four different first languages. Similarly, Vihman (1993) summarizes several longitudinal studies to show that even children who are acquiring the same first language master different phonemes in different orders. Evidence against a clear demarcation between babbling and speech is even more conclusive. Vihman et al. (1985) showed that most children continue to babble for months after they begin to talk. They also showed that the inventory and even the relative frequencies of consonants observed in a child's early words tend to be identical with the inventory and relative frequencies of consonants

produced in the same child's babbling productions just before and during the months that the child is acquiring an initial lexicon of 25-50 words. Therefore, any theory of phonetic and phonological acquisition today must be able to account for the fact that phonological development is highly variable, and that it is continuous with the phonetic patterns that begin to be established in pre-verbal babbling.

2. Laryngeal Features: A Well-Documented Universal

While researchers today readily acknowledge this mounting evidence for variability and continuity, many still accept Jakobson's more general claim that there are substantive universals which structure phonological acquisition. This is because some specific examples of such universals are valid generalizations, particularly if we take the term "universal" to mean a strong numerical tendency rather than an absolute rule and are careful to interpret the example in terms of our better observational techniques today

The best example involves laryngeal features. A search of the UPSID archive of phoneme inventories for 453 languages (Maddieson, 1984) shows that while every language in the database has voiceless unaspirated stops such as /p, t, k/, very few languages contrast them both to voiced stops such as /b, d, g/ and to aspirated stops such as /p^h, t^h, k^h/. These numbers support a generalization about which laryngeal features are "marked" in phoneme inventories across languages, a generalization that accords with Jakobson's claim that "so long as stops in child language are not split according to the behavior of the glottis, they are generally produced as voiceless and unaspirated." Jakobson's claim was based on the few early diary studies available to him, especially Grégoire's (1937) monograph on the acquisition of French. Ingram's (1974a) description of a "pre-vocalic voicing rule" in diary accounts of English-acquiring children such as

Velten (1943) seemed to contradict this claimed universal. However, instrumental studies such as Preston and Kewley-Port (1974) and Macken and Barton (1980a) suggest that when Velten recorded /b, d, g/ for the initial stop in words such as *pie*, *toe*, and *cap*, she was not transcribing true voicing but merely an absence of the aspiration that characterizes English voiceless stops in foot-initial position. See Figure 1 for an example from some of our transcription data. Jakobson's generalization that voiceless unaspirated stops are mastered before either aspirated or voiced stops was also confirmed in instrumental studies of French (Allen, 1985) and several other languages, including Spanish (Macken and Barton, 1980b), Thai (Gandour et al., 1986), Taiwanese (Pan, 1994), and Hindi (Davis, 1995).

*** Insert Figure 1 about here ***

Kewley-Port and Preston (1974) suggest that this universal of acquisition is phonetically grounded in the relative difficulty of satisfying aerodynamic requirements for the different stop types. The build up of oral air pressure during stop closure inhibits voicing even when the vocal cords are adducted, so producing truly voiced stops (i.e., with audible voicing during the oral constriction) requires the child to perform other maneuvers, such as expanding the pharynx. The production of aspirated stops is not so complex, but it does require the child to keep the glottis open exactly long enough after the release of the oral closure to create an audible interval of aspiration during the first part of the following vowel.

In short, the early mastery of voiceless unaspirated stops relative to voiced and aspirated stops seems to be explained well by the inherent motor difficulty of producing the two "marked" stops types. It is thus a good candidate for a phonetic "markedness"

constraint in recent phonological frameworks such as Optimality Theory. That is, it is a well-documented pattern that is so thoroughly grounded in the human capacity for speech production and perception that it seems to be independent of any language-specific facts about the phonological system that the child is acquiring, such as the inventory of contrasts and the relative frequencies of the different stop types in the lexicon of the language. In the rest of this paper, we will focus on two other claims that researchers have assumed to be similarly universal markedness constraints.

3. Lingual Place and Manner Features

The first claim has to do with manner of articulation. Jakobson maintained that stops are universally mastered before fricatives and fricatives before affricates. This claim is supported by the fact that stops seem to occur universally even in babbling, whereas sounds such as /s, ʃ, tʃ/ are among the last consonants to be mastered by children acquiring English (Smit et al., 1990) and French (Chevrie-Muller and Lebreton, 1973). The claim is also supported by substitutions such as those transcribed in the first three sets of examples in Table 1, which are similar to errors observed in other studies of English-acquiring children such as Ingram (1978b).

*** Insert Table 1 about here ***

The second claim has to do with place of articulation. Jakobson maintained that dentals are universally mastered before velars and that "At a particular stage of development ... the Swedish child says *tata* for 'kaka', the German child *topf* for 'kopf', the English child *tut* for 'cut', and the Japanese child also changes *k* to *t*" (p. 46-47). The assumption that "back" consonants such as /k/ are marked and universally liable to be replaced by "front" consonants such as /t/ pervades the literature (e.g., Ingram, 1974,

Locke, 1983, Stemberger and Stoel-Gammon, 1991, Pater, 2003). The fourth and fifth sets of examples in Table 1 illustrate the types of "fronting" that are most frequently observed in English-acquiring children.

The substitutions in the top half of Table 1 are often cited as evidence of acquisition universals, on par with the substitution of voiceless unaspirated "[d]" for English foot-initial aspirated /t/ that is illustrated in Figure 1. However, neither of these claims has been evaluated as thoroughly across languages as the arguments for the unmarked status of voiceless unaspirated stops. When we examine only a single other language (see the substitutions in the bottom half of Table 1), the results suggest that some skepticism is in order. The productions in this part of the table are from a recent study that we describe in Yoneyama, Beckman, and Edwards (2003). The substitution patterns shown here are also observed in earlier studies such as Usizima and Moriwaki (1943) and Nakanishi, Owada and Fujita (1972), and they differ from the patterns most commonly reported for English. For example, the 2-5 year old Japanese-acquiring children who participated in the Yoneyama et al. (2003) study made more than twice as many "backing" errors for /t/ as they made "fronting" errors for /k/. Also, many of the fronting errors were substitutions of [tʃ] or [dʒ] for /k/ rather than of [t] or [d]. By contrast, the 3-5 year old English-acquiring children in Isermann (2001) made three times as many fronting errors for /k/ as they made backing errors for /t/, and only one of these fronting errors was a transcribed substitution of a postalveolar affricate rather than an alveolar stop.

For the sibilant fricatives, similarly, differences in order of acquisition and error pattern are observed between the two languages. For Japanese, /s/ was among the most error-prone sounds for the 4-6 year old children in Nakanishi et al.'s (1972) norming

study. Even a few children in the oldest group (aged 6;6 to 6;11) made errors for target /s/, whereas /ʃ/ was virtually error free in the 5 year olds. Also, backing of /s/ to [ʃ] or [tʃ] is by far the most frequent error that Nakanishi et al. (1972) observed, occurring more than twice as often as fronting of /s/ to [θ]. Fronting of /ʃ/ to [s], [t], or [θ], on the other hand, was extremely rare, occurring less than 10% as often as "stopping" (i.e., the transcribed substitution [tʃ] for /ʃ/). By contrast, in Isermann's (2001) study of English-acquiring children, fronting of /s/ to [θ] and even to [f] occurred twice as often as backing of /s/ to [ʃ], and overall error rate for /s/ was less than a third of the rate of fronting errors for /ʃ/.

If stops are universally unmarked relative to affricates and if front consonants are universally unmarked relative to back, how can we explain the substitution patterns that Japanese-acquiring children show? Could there be characteristic phonetic properties of some of these Japanese sounds that are obscured in the transcription, properties comparable to the aspiration that distinguishes English /p, t, k/ in word-initial position even from voiceless variants of /b, d, g/? Alternatively, could there be some higher-level difference between the phonological systems of Japanese and English (and French), something more subtle than the presence or absence of particular contrasts, that might explain why Japanese-acquiring children back /t/ to [k] and /s/ to [ʃ]? We turn to these questions in the next three sections.

4. Phoneme Frequency Effects

As suggested in Section 2, the term "marked" in phonology refers to a complex of several properties, which in the case of laryngeal features happen to be correlated. One

of those properties is the relative phonetic “complexity” or articulatory difficulty associated with the phonological features being compared; word-initial voiced stops are “marked” relative to voiceless stops because initiating voicing during a stop closure requires a simultaneous expansion of the oral cavity to maintain an adequate pressure differential across the glottis. Another property is the “typicality” or relative frequency of occurrence across different languages; although all languages have voiceless unaspirated stops, not all languages contrast them to voiced stops.

If we apply this frequency measure of typicality to account for the substitution of [t] or [k] for /tʃ/ in the second set of examples in Table 1, we can say that affricates are marked relative to stops, since many of the languages in the UPSID archive have no affricates at all, whereas all of them have at least three stops. However, relative typicality by this measure cannot account for the “canonical” stopping illustrated in the first set of examples in Table 1, the substitution of [t] for /s/ and so on. Nor can it cannot for the fronting of /k/ to [t] in the fourth set of examples, since each of the sounds /s/, /k/, and /t/ occurs in the phoneme inventories of nearly all of the languages in the UPSID archive, making them all equally typical or unmarked.

Maddieson (1997) gives an alternative measure of typicality — the number of words “within the lexicon of particular languages” (p. 637), a measure which he finds to be in good, but not perfect correlation with the number of languages in which a sound occurs. While /k/ is no more marked than /t/ by the cross-language frequency measure, the two sounds might be distinguished by the lexicon-based measure, which has two advantages over the inventory-based measure. First, since different languages have different lexicons, a lexicon-based frequency measure can in theory yield different markedness

rankings in different languages. For example, Pye et al. (1987) suggest that the reason that Quiché-acquiring children master /tʃ/ at an earlier age than English-acquiring children is because this consonant occurs in many words of Quiché but is the least frequent of the eight plosives of English. Second, since acquiring language means learning the words and grammar of a specific language, it is much easier to see how the language-specific measure might be incorporated into an explanatory model of acquisition. For example, Stemberger and Bernhardt (1998) suggest that fronting is a typical substitution pattern for many young children because /t/ occurs in more words than /k/ in the lexicons of English, German, and many other languages. For children acquiring these languages, perception and production of /t/ should be more practiced and automatic, requiring fewer mental resources than the perception and production of /k/.

In work currently in progress, we are evaluating the usefulness of this measure of language-specific markedness by analyzing speech addressed to very young children for several different languages, and tabulating the relative frequencies of each of the lingual obstruent phonemes in the words that the children hear addressed to them. (We are analyzing these "child-directed vocabularies" rather than the available adult lexicons because we cannot be sure that the relative frequencies for the target sounds are the same in the two speaking styles.) Results to date show that, although prevocalic /k/ is somewhat more frequent than /t/ in the online English dictionary that we have consulted (*Hoosier Mental Lexicon*, Pisoni et al., 1985), /t/ occurs more often than /k/ in the words that an English-acquiring child is likely to learn first. In Japanese, by contrast, /k/ occurs far more frequently than /t/ both in the adult lexicon (NTT database, Amano and Kondo, 1999) and in the words that a Japanese-acquiring infant is likely to hear (Yoneyama et al.,

2003). This higher type frequency for /k/ relative to /t/ in Japanese predicts the earlier acquisition of /k/ reported in Usizima and Moriwaki (1943) and the larger number of errors for /t/ reported in Yoneyama et al. (2003).

While we have not yet tabulated the sibilants in the child-directed speech, our counts of words in adult lexicons suggest that different relative frequencies for [s] versus [ʃ] in the two languages will also predict the language-specific pattern of fronting of /ʃ/ to [s] in English-acquiring children, as contrasted to the backing of /s/ to [ʃ] in Japanese-acquiring children. That is, /s/ is very frequent in the adult lexicons of both languages, but its frequency relative to /ʃ/ is much greater in English, with a ratio of more than 6 to 1 in word-initial position, as compared to only 1.2 to 1 in Japanese. Of course, this measure of language-specific markedness predicts more /ʃ/ to [s] fronting errors in English, but it cannot predict the prevalence of /s/ to [ʃ] backing errors in Japanese except in combination with some other measure of markedness, which we will describe in Section 5. It also cannot explain the alveolar fronting errors that are frequently observed in English-acquiring children, since /s/ occurs in 17 times as many words as /θ/.

Moreover, while language-specific markedness as measured by the relative type frequencies of /t/ versus /k/ and /s/ versus /ʃ/ does seem to predict some of the fronting or backing substitutions illustrated in the two parts of Table 1, it does not seem likely that the relative frequencies of /s/ versus /t/ in English or of /s/ versus /tʃ/ in Japanese can predict the two patterns of “stopping” of fricatives, or the very late acquisition of /s/ and

/ʃ/ relative to stops in both languages (see, e.g., Smit et al., 1990, for English, and Nakanishi et al., 1972, for Japanese).

We think that the late acquisition of the strident fricatives relative to the two plosives might be explained by the extreme articulatory precision required to maintain strong airflow and direct the resulting stream of air against the incisors. That is, [s] and [ʃ] are more difficult to produce than [t] in the same way that truly voiced stops are more difficult to produce than voiceless unaspirated stops.[1)] Of course, proposing such an explanation is risky, because there are no standard measures of this kind of markedness comparable to the two frequency measures described in this section. We have already suggested that the earlier acquisition of voiceless unaspirated stops relative to prevoiced and voiceless aspirated stops is due to the more difficult articulatory manuevers required for these latter two types of stops. In the next section, we describe a phenomenon that has been cited as evidence for "articulatory markedness" in the literature on laryngeal contrasts and that can be extended to other cases of articulatory markedness as well.

5. Covert Contrast

Cataloguing error patterns in transcribed forms like those in Table 1 is the predominate method used in studying phonological acquisition. Macken (1980) suggests that there are at least two different ways to interpret these apparent "substitutions" and these interpretations cannot be distinguished in the standard cross-sectional study such as Nakanishi et al. (1972) or Smit et al. (1990). She calls these interpretations "perceptually based rules" versus "articulatorily based rules" and they reflect two different hypotheses about the child's "underlying representations" — i.e., the internal representations of words in the young child's rapidly expanding vocabulary.

The apparent substitutions are "perceptually-based" when the child does not have a reliable internal representation of the distribution of acoustic patterns that contrast the target adult sound with the "substituted" sound. For example, the English-acquiring child who is transcribed as producing [d] for /t/ in *tub* might perceptually categorize the onset together with the onset of the word *dub*. In this case, the child has only one category, which is identical with (or at least closer to) the adult's /d/. The substitutions are "articulatorily-based" when the child has a reliable internal representation of the acoustic differences between the two segments, but has not mastered the motor representations necessary to producing the differences reliably. For example, the English-acquiring child who is transcribed as producing "[d]" for /t/ might have robust perceptual representations of *tub* versus *dub*, but not have acquired the motor control to make the peak of the glottal opening gesture occur at the right time relative to the release of the alveolar closure.

These two interpretations cannot be differentiated simply by looking at the transcriptions, but Macken and Ferguson (1981) suggest that "articulatorily-based rules" (and "articulatory markedness") can be identified with the phenomenon of "covert contrast" demonstrated in Macken and Barton's (1980a) longitudinal study of the acquisition of voicing contrasts in English. If there is an established acoustic measure that correlates well with the contrast between the substituted sound and the target, we can define covert contrast as a statistically significant acoustic difference in that measure that is not large enough to be perceived reliably by the transcriber. In addition to transcribing each children's productions, Macken and Barton also measured the voice onset time (VOT) of the target voiceless stops. In the earliest recording sessions, most of the children produced VOT values at or near 0 seconds both for target /b, d, g/ and for target

/p, t, k/, just as in the initial /b/ in *bath* in both productions in Figure 1 and the initial /t/ in *tub* in p174's production in that figure. Many of the children also went through a stage where they produced significantly longer VOT values for /p, t, k/, although not long enough for these to be perceived reliably as /p, t, k/ by the transcriber

Covert contrast has been demonstrated for some of the characteristic place substitutions described in Section 3. For example, White (2001) used several acoustic measures of the distribution of energy in the spectra of bursts of target /t/ and /k/ produced by fourteen English-acquiring children who had atypical pronunciations of these sounds in familiar words. That is, these children were transcribed as substituting [t] for /k/ or [k] for /t/ on the /k/ and /t/ targets either in words elicited in a standardized test of articulation (GFTA, Goldman and Fristoe, 1986) or in words elicited in Isermann's (2001) study. Figure 2 shows sample spectra in one of the children in the control group of fourteen children who did not front /k/ or back /t/. For the stops before /i/, White calculated the mean frequency and skewness, as in Forrest et al. (1990). For the stops before /u/, he measured the frequency of the lowest-frequency peak in the LPC-smoothed spectrum and the proportion of energy centered around this peak. Figure 3 shows representative results for the last two measures. The child in panel (a) produced a typical contrast between the two stops; tokens of /ku/ are well separated from tokens of /tu/, with a more compact (higher proportion of energy) and lower frequency peak. The two children in the other panels were both transcribed as producing [k] for /t/ in this back vowel context, but their acoustic patterns differ. The child in panel (b) shows covert contrast; there is a statistically significant separation between the centers of the data clouds, but many tokens of /tu/ have somewhat compact peaks and/or somewhat low

peaks, making the two data clouds overlap. The child in panel (c), on the other hand, shows something more like Macken's "perceptually based rule" — a pattern which White (2001) terms a "merger"; both data clouds are centered in an intermediate region, suggesting an "undifferentiated lingual gesture" (see also Gibbon, 1999).

*** Insert Figures 2 and 3 about here ***

Baum and McNutt (1990) similarly studied fronted productions of /s/ as [θ] in English-acquiring children, using such measures as the amplitude of the fricative relative to the following vowel, which should be higher in the sibilant fricative /s/ than in the non-strident /θ/. Their results show that productions of target /s/ that are transcribed as [θ] are significantly different from productions of target /θ/, although the differences are not large enough to cue the percept of /s/. Tsurutani (2002) studied backed productions of /s/ as [ʃ] in Japanese-acquiring children using two spectral measures: the "centroid" value in the fricative (i.e., mean frequency of a spectral slice taken at the middle of the fricative), which should be higher in /s/ than in /ʃ/, and second formant value at the onset of the following vowel, which should be lower in /s/. She found a small but reliable difference between productions of target /s/ that were transcribed as [ʃ] and productions of target /ʃ/.

These two studies of covert contrast in the typical patterns of error for /s/ in the two languages suggest that /s/ has more difficult articulatory requirements than /θ/ or /ʃ/. Perhaps there is only a narrow range of lingual postures and constriction locations that will reliably produce the percept of /s/, and children acquiring a language that contrasts /s/ to some other laminal or apical fricative might initially miss that range in one direction

or the other. Whether the typical direction of the "miss" is in the direction of the more back sibilant, as in Japanese, or toward a more front articulation, as in English, might then depend on other language-specific facts besides the relative frequencies of /s/ versus /ʃ/ (or /s/ versus /θ/) in the lexicon.

One potentially relevant fact is subtle differences in the articulations of the sounds between the two languages. In particular, /ʃ/ in English has a lower centroid frequency due to the presence of lip rounding. Two more potentially relevant facts are the presence in the phoneme inventory of other lingual fricatives (/θ/ in English versus /ç/ in Japanese) and the different sequential constraints in the two languages. In particular, whereas /s/ combines readily with following /i/ and /e/ in many words of English, /s/ does not occur before /i/ in any word of Japanese, and /e/ itself is very infrequent, giving the Japanese-acquiring child considerably less experience with /s/ before front vowels. A further fact of note is that there are nearly three times as many words beginning with a lingual obstruent followed by a front vowel as a back vowel in English, whereas there are many more words beginning with a lingual obstruent followed a back vowel in Japanese. These kinds of language-specific facts might conspire to induce a difference in "basis of articulation" (Heffner, 1950) between the two languages — i.e., a difference in the "characteristic posturing of tongue and lips" which Vihman (1996, p. 120) suggests might begin to be set in place already in infancy, from visual cues.

6. Consonant-Vowel Interactions

In the previous section, we speculated that the relatively low type frequency of /e/ in Japanese interacting with the sequential constraint against /si/ might contribute to a more "palatal" articulatory basis for sibilants in the language. This idea was suggested to us by

certain other results of the Yoneyama et al. (2003) study, which we summarize briefly here. In that study, the overall error rate for /t/ was much higher than the error rate for /k/. However, one context was an exception. Children fronted /k/ before /e/ as often as they backed /t/ there. Also, the substitution pattern was different in this context. Where dentals were the more frequent sound to substitute for /k/ in other contexts and velars were frequently substituted for /t/ before /a/ and /o/ (the other two vowel contexts we tested), both /k/ and /t/ were frequently transcribed as [tʃ] in /ke/ and /te/. Nakanishi et al. (1972) do not differentiate /ke/ from /ka/, /ku/, and /ko/ in their norming study, so it is impossible to know how many of the [tʃ] for /k/ errors that they counted were produced before /e/. However, they do provide separate counts for /ki/, and there, fronting to [tʃ] occurs three times as often as fronting to [t]. By contrast, English-acquiring children almost never front to /tʃ/ in any environment, as we noted above in Section 3. Again, we wonder whether these differences between the two languages might be related to the very low type frequency of /e/ in Japanese interacting with the constraint against /ti/, which is an extremely marginal CV sequence that has only come into the language in recent loan words such as ティッシュ.

White (2001) shows a related context effect for lingual stops in English-acquiring children. The two children who backed /t/ to [k] in Figure 3 were exceptional. Most of the children with atypical productions in this study had the opposite substitution, fronting /k/ to [t], as is typical in English. However, they fronted more often in the /i/ context than in the /u/ context, and there was one child who fronted /ki/ to [ti] but backed /tu/ to [ku]. The /ku/ versus /ki/ burst spectra in Figure 2 suggest why there might be such effects of

vowel context on the substitution patterns for /k/. The velar stops in languages like English and French have more anterior contact in the context of front vowels. Spectral energy is concentrated in a higher frequency region, making the /ki/ burst similar to the /ti/ burst.

Keating and Lahiri (1993) equate these "fronted velars" of English and French with the contrastively palatalized velars of languages such as Russian. In Japanese, the contact of /k/ in /ki/ is particularly front, because of the very close and front articulation of the /i/ itself (see Wada et al. 1969). There is a longer release burst in this context as well, because of the slower venting of air pressure behind the narrow constriction in the /i/. This makes the substitution of [tʃ] for /k/ in the context of /i/ in the Japanese-acquiring children observed in Nakanishi et al. considerably less surprising.

6. Summary

In this paper, we have reviewed some older and more recent results on the acquisition of lingual obstruents in Japanese-acquiring children. Comparing these results to studies of English-acquiring children reveal some commonalities and some initially surprising differences between the two languages. We have speculated on the reasons for these apparent differences. These speculations, we hope, have also given a flavor of the kind of research that we think will be necessary in order to understand what it means for a developmental pattern to be "universal".

Acknowledgments

This work was supported in part by NIDCD grant 02932 to Jan Edwards and by an Ohio State University Cognitive Science Summer Fellowship to Kiyoko Yoneyama.

Notes

1) The fact that /s/ occurs in most phoneme inventories and in many words in the lexicon despite this “articulatory markedness” might then be due to its relatively “unmarked” status by some other criterion, such as perceptual salience or “ease of contrast”. The high-amplitude turbulent energy that results from directing a stream of air against the incisors makes [s] (and [ʃ]) stand out from neighboring vowels, and also makes the two sibilants contrast well both with other continuant sounds such as [f] or [h] and with the voiceless unaspirated stops that are the most common stop type cross-linguistically.

References

Allen, G. D. (1985). “How the young French child avoids the pre-voicing problem for word-initial voiced stops,” *Journal of Child Language*, 12, 37-46.

Amano, S. and T. Kondo (1999) *Lexical Properties of Japanese*. Sanseido. （天野成昭・近藤公久） (1999) 『日本語の語彙特性』（「NTT データベースシリーズ」）東京：三省堂）

Baum, S. R. and J. C. McNutt (1990) “An Acoustic Analysis of Frontal Misarticulation of /s/ in Children,” *Journal of Phonetics* 18, 51-63.

Chevrie-Muller, C. and M. T. Lebreton (1973) “Étude de la réalization des consonnes au cours d’une épreuve de mots, sur des groupes d’enfants de 3 ans et 5 ans ½,” *Revue de laryngologie* 94, 109-152.

Davis, K (1995) “Phonetic and Phonological Contrasts in the Acquisition of Voicing: Voice Onset Time Production in Hindi and English,” *Journal of Child Language* 22, 275-305.

Forrest, K., G. Weismer, M. N. Hodge, D. A. Dinnsen, and M. Elbert (1990) “Statistical Analysis of Word-Initial /k/ and /t/ Produced by Normal and Phonologically Disordered Children,” *Clinical Linguistics and Phonetics* 4, 327-340.

Gandour, J., H., S. H. Petty, R. Dardarananda, S. Dechongkit, and S. Mukongoen (1986) “The acquisition of the Voicing Contrast in Thai: A Study of Voice Onset Time in Word-Initial Stop Consonants,” *Journal of Child Language* 13, 561-572.

Gibbon, F. E. (1999) “Undifferentiated Lingual Gestures in Children with Articulation/Phonological Disorders,” *Journal of Speech, Language, and Hearing Research* 42, 382-397.

Goldman, R. and M. Fristoe (1986) *The Goldman Fristoe Test of Articulation*. Circle Pines, MN: American Guidance Services.

Grégoire, A. (1937) *L’apprentissage du langage*. Bibliothèque de la Faculté de Philosophie et Lettres de L’Université de Liège, 73.

Heffner, R. M. S. (1950) *General Phonetics*. Madison: University of Wisconsin Press.

Ingram, D. (1974a) “Phonological Rules in Young Children,” *Journal of Child Language* 1, 49-64

Ingram, D. (1974b) “Fronting in Child Phonology,” *Journal of Child Language* 1, 233-241.

Ingram, D. (1978) “The Acquisition of Word-Initial Fricatives and Affricates by Normal and Linguistically Deviant Children.” In A. Caramazza and E. B. Zurif (eds.) *Language Acquisition and Language Breakdown: Parallels and Divergences.* Baltimore: Johns Hopkins University Press.

Ingram, D. (1999) “Phonological acquisition.” In M. Barrett (ed.) *The Development of Language* (pp. 73-97). Hove, UK: Psychology Press.

Isermann, B. C. (2001) *Variability of Consistency of Articulation in Children with Phonological Disorders*. M.A. thesis, Ohio State University.

Jakobson, R. (1941/1968) *Kindersprache, Aphasie und allgemeine Lautgestze*. [Published originally in 1941; citations are from the 1968 translation by A. R. Keiler, *Child Language, Aphasia, and Phonological Universals*. The Hague: Mouton.]

Keating, P. and A. Lahiri (1993) “Fronted Velars, Palatalized Velars and Palatals,” *Phonetica* 73-101.

Kewley-Port, D. and M. S. Preston (1974) “Early Apical Stop Production: A Voice Onset Time Analysis,” *Journal of Phonetics* 2, 195-210.

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

Macken, M. A. (1980) “The Child’s Lexical Representation: The ‘puzzle-puddle-pickle’ Evidence,” *Journal of Linguistics* 16, 1-17.

Macken, M. A. and D. Barton (1980a) “The Acquisition of the Voicing Contrast in English: A Study of Voice Onset Time in Word-Initial Stop Consonants,” *Journal of Child Language* 7, 41-74.

Macken, M. A. and D. Barton (1980b) “The Acquisition of the Voicing Contrast in Spanish: A Phonetic and Phonological Study of Word-Initial Stop Consonants,” *Journal of Child Language* 7, 433-458.

Macken, M. A. and C. A. Ferguson (1981) “Phonological Universals in Language Acquisition.” In H. Winitz (Ed.) *Native Language and Foreign Language Acquisition* (pp. 110-129). New York: New York Academy of Sciences.

Maddieson, I. (1984) *Patterns of Sounds*. Cambridge: Cambridge University Press.

Maddieson, I. (1997) “Phonetic Universals.” In W. J. Hardcastle and J. Laver (Eds.) *The Handbook of Phonetic Sciences* (pp. 619-639). Oxford: Blackwell.

Nakanishi, Y., K. Owada, and N. Fuzita（中西靖子・大和田健次郎・藤田紀子）(1972)「構音検査とその結果に関する考察」『特殊教育研究施設報告』（東京学芸大学特殊教育研究施設編）1, 1-41.

Pater, J. (2003) “Form and Substance in Phonological Development.” In L. Mikkelsen and C. Potts (eds.) *Proceedings of the 21st West Coast Conference on Formal Linguistics*. (pp. 348-372.) Somerville, Mass.: Cascadilla Press.

Pan, H. (1994) *The Voicing Contrasts of Taiwanese (Amoy) Initial Stops: Data from Adults and Children*. Ph. D. thesis, Ohio State University.

Pisoni, D., H. Nusbaum, P. Luce and L. Slowiacek (1985) “Speech Perception, Word Recognition, and the Structure of the Lexicon,” *Speech Communication* 4, 75-95.

Pye, C., D. Ingram and H. List (1987) “A Comparison of Initial Consonant Acquisition in English and Quiché.” In K. E. Nelson and A. Van Kleeck (eds.) *Children’s Language, 6* (pp. 175-190.) Hillsdale, NJ: Lawrence Erlbaum.

Smit, A. B., L. Hand, J. J. Feilinger, J. E. Bernthal and A. Bird (1990) “The Iowa Articulation Norms Project and its Nebraska Replication,” *Journal of Speech and Hearing Disorders* 55, 779-798.

Stemberger, J. P. and C. Stoel-Gammon (1991) “The Underspecification of Coronals: Evidence from Language Acquisition and Performance Errors.” In C. Paradis and J-F Prunet (eds.) *The Special Status of Coronals: Internal and External Evidence*. (pp. 181-199). San Diego: Academic Press.

Stemberger, J. P. and B. H. Bernhardt (1998) “The Emergence of Faithfulness.” In B. MacWhinney (ed.) *The Emergence of Language* (pp. 417-446.) Mahwah, NJ: Lawrence Erlbaum.

Tsurutani, C. (2002) “Acquisition of Japanese Contracted Sounds in L1 Phonology,” Poster presented at the 8th Conference on Laboratory Phonology, Yale University, 27-30 June 2002.

Usizima, Y. and K. Moriwaki（牛島良知・森脇要）(1943) 「幼児の文章の構造」『幼児の言語発達』（恩賜財団愛育会愛育研究所編）55-79、東京：目黒書店.

Velten, H. (1943) “The Growth of Phonemic and Lexical Patterns in Infant Language,” *Language* 19, 440-444.

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

Vihman, M. M. (1996) *Phonological Development: The Origins of Language in the Child*. Oxford: Blackwell.

Vihman, M. M., M. A. Macken, R. Miller, H. Simmons and J. Miller (1985) “From Babbling to Speech: A Re-Assessment of the Continuity Issue,” *Language* 61, 397-445.

Wada, T., M.Yasumoto, N. Ikeoka, Y. Fujiki and R. Yoshinaga (1969) “An Approach for the Cineflourographic Study of Articulatory Movements,” *The Cleft Palate Journal* 7, 506-522.

White, S. D. (2001) *Covert Contrast, Merger, and Substitution in Children’s Productions of /k/ and /t/.* M.A. thesis, Ohio State University.

Yoneyama, K., M. E. Beckman and J. Edwards (2003) “Phoneme Frequencies and Acquisition of Lingual Stops in Japanese.” Manuscript under review.

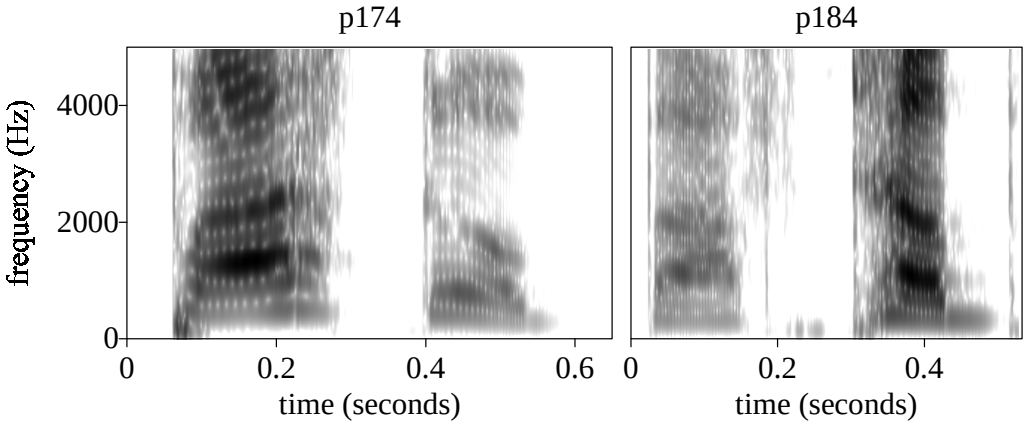

Figure 1. Spectrograms of the word *bathtub* transcribed (left panel) as [bæθ dʌb] and (right panel) as [bæθ tʌb] in Isermann's (2001) study. Child No. p174 who produced the apparent /d/ (voiceless unaspirated [t]) in *tub* is a male aged 4(years);10(months). Child No. p184 who produced the apparent /t/ (aspirated [t^h]) is a male, aged 4;9. The initial /b/ is voiceless unaspirated [p] in both productions, as is typical in this dialect.

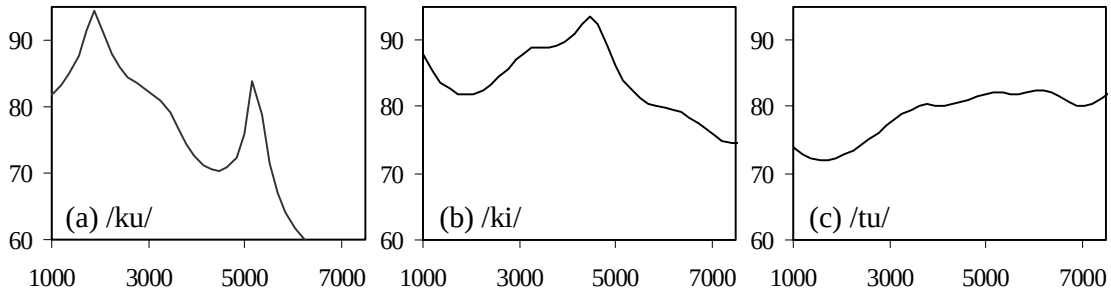

Figure 2. Spectra (amplitude in dB against frequency in Hz) calculated over 10 ms windows centered at the stop bursts for representative tokens by participant n011 in White (2001).

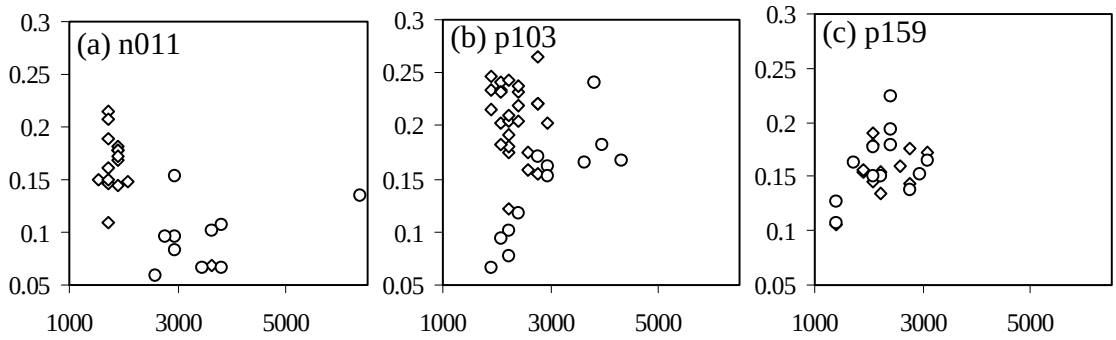

Figure 3. Proportion of spectral energy in a 700 Hz band around the lowest frequency peak in the spectrum plotted against the frequency (Hz) of that first peak for (a) a child with a typical /k/-/t/ contrast, (b) a child showing covert contrast, and (c) a child showing merger in a region between /k/ and /t/. The circles represent tokens of /ku/ and the diamonds represent tokens of /tu/.

Table 1. Substitutions of place and manner in two studies

"phonological process"	target	adult form	child form	ID sex age (yr; mo)
a. English-acquiring children with phonological disorder (Isermann, 2001)				
stopping — substitution of stops for fricatives	*socks*	/sɑks/	[dath]	p137 M 4;4
	sheep	/ʃip/	[ti]	p112 F 5;4
"stopping" — substitution of stops for affricates	*chair*	/tʃɛɹ/	[teo]	p116 M 4;7
	cheese	/tʃiz/	[ti]	p105 M 4;1
	cheese	/tʃiz/	[ki]	p103 F 5;9
simplification of affricates to fricatives	*cherries*	/tʃɛɹiz/	/ʃɛwiz/	p105
	chicken	/tʃɪkən/	[θɪkɪn]	p121 M 3;11
fronting of velars	*cake*	/keɪ̯k/	[teɪ̯k]	p106 F 5;7
	goat	/got/	[dot]	p106
fronting of postalveolars	*brush*	/bɹʌʃ/	[bwʌs]	p106
	shoe	/ʃu/	[su]	p124 M 4;11
b. Japanese-acquiring children with typical development (Yoneyama et al., 2003)				
"stopping" — substitution of affricates for fricatives	猿	/saɾu/	[tʃaɾu]	t20 F 3;8
	障子	/ʃo:dʒi/	[tʃo:tʃi]	t21 F 3;9
assimilation — substitution of affricates for stops	テープ	/te:pu/	[tse:pu]	t10 M 2;7
	ケーキ	/ke:ki/	[tʃe:ki]	t10
backing of dentals	狸	/tanuki/	[kik:i]	t04 M 2;4
	虎	/toɾa/	[koɾa]	t17 F 3;6
backing of alveolars	掃除機	/so:dʒiki/	[ʃo:dʒik:i]	t17
	シャツ	/ʃatsu/	[ʃatʃu]	t22 F 3;9
	猿	/saɾu/	[ʃaɾu]	t22