Generalizing over lexicons to predict consonant mastery

MARY E. BECKMAN and JAN EDWARDS

Ohio State University
University of Wisconsin – Madison

Abstract

When they first begin to talk, children show characteristic consonant errors, which are often described in terms that recall Neogrammarian sound change. For example, a Japanese child's production of the word kimono *might be transcribed with an initial postalveolar affricate, as in typical velar-softening sound changes. Broad-stroke reviews of errors list striking commonalities across children acquiring different languages, whereas quantitative studies reveal enormous variability across children, some of which seems related to differences in consonant frequencies across different lexicons. This paper asks whether the appearance of commonalities across children acquiring different languages might be reconciled with the observed variability by referring to the ways in which sound change might affect frequencies in the lexicon. Correlational analyses were used to assess relationships between consonant accuracy in a database of recordings of toddlers acquiring Cantonese, English, Greek, or Japanese and two measures of consonant frequency: one specific to the lexicon being acquired, the other an average frequency calculated for the other three languages. Results showed generally positive trends, although the strength of the trends differed across measures and across languages. Many outliers in plots depicting the relationships suggested historical contingencies that have conspired to make for unexpected paths, much as in biological evolution.*

> "The history of life is not necessarily progressive; it is certainly not predictable. The earth's creatures have evolved through a series of contingent and fortuitous events." (Gould 1989)

1. Introduction

When children first begin to produce vocalizations that listeners recognize as meaningful words of the ambient language, they show characteristic consonant misarticulations. These errors are often perceived in terms of categorical processes that substitute one target for another. For example, Cantonese-learning children are perceived as substituting unaspirated stops for aspirated and velars for labialized velars, so that a 2-year-old's produc-

tions of the words [kʰok⁵⁵kʰei²¹pɛːŋ³⁵] 'cookie' and [kʷaːi³³sɐu³³] 'monster' might both be transcribed as having initial [k].[1] Similarly, Japanese children are perceived as fronting contrastively palatalized dorsal stops and even allophonically palatalized ones, so that a 2- or 3-year-olds' productions of *kyōsō* 'race' might be transcribed with an initial [tʃ] and so too might be a 2-year-old's productions of *kimono* 'kimono' and *kemuri* 'smoke'.

There are two seemingly contradictory characterizations of these errors. First, work such as Locke's (1983) masterful review monograph offers broad-stroke generalizations about typical consonant inventories of very young children acquiring different languages, generalizations that tend to support Jakobson's (1941/1968) "laws of irreversible solidarity" — i.e., implicational universals such that mastery of "complex," "elaborated," or "marked" consonants implies prior mastery of "simple," "basic," or "un-marked" ones.[2] By contrast, quantitative observations of individual tokens in more controlled cross-language studies such as Vihman (1993) often show enormous variability in production patterns, at least part of which seems to be related to differences in consonant frequencies across the dif-ferent lexicons that children can acquire (e.g., Pye, Ingram and List 1987).

Research in the first vein draws parallels between universals in acquisi-tion and the markedness hierarchies noted for consonant inventories across languages and for patterns often seen in loanword adaptation and sound change. For example, [k] is listed in 403 of the 451 UPSID languages (Maddieson and Precoda 1989) whereas [kʷ] and contrastively palatalized [kʲ] are listed in only 60 and 13, respectively. Also, the Cantonese-learning child's substitution of [k] for [kʷ] mirrors the delabialization of this initial in cognate forms such as Sino-Japanese *kaidan* 'ghost story' as well as the delabialization of Latin [kʷ] in modern French forms such as *quotient* ([kɔsʲã]) and *que* ([kʲe]). The Japanese-learning child's substitution of [tʃ] for [kʲ] likewise mirrors the velar softening processes that resulted in post-alveolar affricate reflexes for earlier dorsal initials in Italian *cielo* 'sky', English *cheese*, Putonghua [tɕi⁵⁵] 'chicken', and so on. Moreover, just as younger Japanese toddlers are more likely than older ones to show this substitution for allophonically palatalized stops as well as for contrastively palatalized ones, comparison across languages in which velar softening is attested shows a hierarchy of targeted contexts such that application of the change to dorsal stops in the context of [e] or [i] implies its application to contrastively palatalized dorsal stops (Guion 1998).

Research in the second vein, on the other hand, emphasizes "the pri-macy of lexical learning in phonological development" (Ferguson and Far-well 1975: 36), noting that children do not master consonants such as [kʷ], [kʲ], and [tʃ] in isolation. Indeed, many consonants cannot be produced audibly in isolation. The sound of [kʷ] or [kʲ] when there is no neighboring vowel or vowel-like release is a silence indistinguishable from the sound of

the closure phase of [k] or [tʃ]. Mastery of a consonant implies mastery of a set of words containing the consonant in a particular set of contexts that allow the consonant to be audibly pronounced for the child to hear and reproduce. Learning many words with a particular sound pattern gives the child practice in resolving the mapping between acoustics and articulation, leading to more robust abstraction of the consonant away from the particular word context (cf. Beckman, Munson and Edwards 2007, *inter alia*).

In this paper, we examine whether the two characterizations might be reconciled by referring to different paths of convergence in the distribution of consonants across the lexicons of children and across the lexicons of languages over time. In identifying these paths, we are mindful of Vygotsky's (1978) view of child development as a process that takes place in environments that are shaped by cultural evolution as well as by biological evolution. As Kirby, Dowman and Griffith (2007) note, this means that explanations for language universals must refer to potential interactions among adaptive systems operating at three different time scales, to shape the course of individual development, the course of cultural transmission, and the course of biological evolution. Thus, to understand the implicational universals that are identified in broad-stroke characterizations of children's production errors, we must disentangle two types of potential influence from pan-species capacities.

The first type of influence is direct. A marked consonant could be difficult to master and an unmarked one easy to master because the child's capacity for speech production and perception is constrained by the same biological and physical factors that are at play at the beginning stages of common sound changes. For example, a palatalized dorsal stop before [i] has a long release interval for the same aerodynamic reasons that stops generally have longer voice onset time values before high vowels. This long VOT interval also has a high-frequency concentration of energy, so that the release phase of [kʲi] is inherently confusable with the strident frication phase of [tɕi], making velar softening a prototype example of listener-based sound change (Ohala 1992, Guion 1998). This ambiguity also should make the sequence difficult for a Japanese child to parse correctly in the input, leading to a lower accuracy rate for the child's reproduction of [kʲi] relative to [ka].

The second type of influence is indirect. A sound change that is observed in many languages could result in cross-language similarities in consonant frequency relationships, because the marked consonant that is the target of the change will be attested in fewer words and the unmarked consonant that is the output of the change will be attested in more words after the change has spread through the speech community. For example, while Latin had allophonically fronted [kʲ] in many words, including the word that developed into *cielo*, modern Italian has this sound only in a few

forms such as *chianti*. The comparable unaspirated ("voiced") dorsal stop of Modern English has a similar distribution, occurring before a high front vowel only in *geese*, *giggle*, *guitar*, and a handful of other forms that might be said to a young child. A consonant that is rare in the input because of this kind of cultural evolution of the lexicon should have lower accuracy rates because there is less opportunity to abstract knowledge of its sound pattern away from known words in order to learn new words.

To begin to disentangle the direct path of influence from the indirect path of influence, we did two sets of correlational analysis involving the target consonants in a database of word productions elicited from children acquiring Cantonese, English, Greek, or Japanese. First, we compared transcribed accuracy rates and relative type frequencies for shared target consonants across each language pair, to see whether patterns of variability in the accuracy rates were correlated between the two languages, and if they were, whether the correlations could be explained by correlations in relative frequencies in the two lexicons. Second, we calculated four sets of mean frequencies for all consonants attested in any of the four lexicons, with each set excluding the frequencies in one of the target languages, to make a more general measure of relative markedness that is not based on frequencies in that particular target language. We then used this language-general measure as the independent variable in a set of regression analyses with language-specific frequencies or accuracy rates as the dependent variable, in order to assess the degree to which the language-general measure of average relative markedness could predict the accuracy patterns for each individual language, once the contribution of type frequencies in the target language is partialled out.

In the following sections, we describe the target consonant productions, motivate the equivalence classes that we set up to compare accuracies and frequencies across the languages, and report the results of the two sets of analyses. In general, we found that the second type of analysis was more informative than the first, although the predictive power of the regression functions was highly variable across the languages. At the same time, patterns in the residuals also were informative when interpreted in terms of the language-specific history of sound changes and circumstances of language contact. Our conclusion from this exercise will be that the relationship between the evolution of consonant systems in language change and the development of consonant systems in acquisition, like the relationship between biological evolution and ontogeny of species, is complicated by the effects of singular events — i.e., of what Gould (1989) calls "the contingency of evolutionary history."

2.   The target consonant productions

The accuracy data are native-speaker transcriptions of target consonants in word productions recorded for the παιδολογος project, a large on-going cross-sectional, cross-linguistic study of phonological acquisition. To date, we have recordings of 20 adults and at least 100 children for each of Cantonese, English, Greek, Japanese, Korean, and Mandarin. The project and transcription methods are described in more detail in Edwards and Beckman (2008a, 2008b). Summarizing briefly: we elicit productions of a word by presenting participants with a picture prompt and a recorded audio prompt that was produced in a child-directed speaking style by an adult female native speaker of the language, and asking them to name the picture by repeating the audio prompt; productions are digitally recorded, and the target consonant and following vowel are transcribed by a native speaker who is a trained phonetician in a two-stage process whereby the consonant (and following vowel) are first identified as "correct" or "incorrect" and then analyzed further to note the nature of the error for incorrect production; finally, transcription error is calibrated by having a second native speaker independently transcribe 10% of the data in the same way.

In this paper, we report data from the first phase of the project, which included about 10 two-year-old children and 10 three-year-old children for each of the first four languages, with inter-transcriber agreement rates for "correct" versus "incorrect" ranging between 90% and 96% by language. The children included roughly equal numbers of boys and girls for each age group for each language, but the sampling across the months for the two-year-old group was more even for the Cantonese- and English-acquiring children than for the Greek- and Japanese-acquiring children.

In these first-phase recordings, the target consonants for each language were all of the lingual obstruents that occur word-initially in pre-vocalic position in at least three words that a child is likely to know. Also, whenever possible, we elicited a consonant in three real words for each of five types of following vowel environment where the consonant is phonotactically possible in word-initial position in a particular target language.[3]

Table 1 lists the 29 word-initial lingual obstruents that we elicited in any of the four languages, along with example words. If there were enough familiar words exemplifying a particular consonant type in a particular language to justify eliciting that target from the two- and three-year-old children in the study, the example word is one that we used in the word repetition task. Otherwise, the example word is enclosed in parentheses.

Table 1. Plotting symbols and example words for the 29 word-initial consonant types elicited in any of the four languages, ordered by mean frequency across lexicons.

IPA and WorldBet		Example word if attested (in parentheses if target C not elicited)			
		Cantonese	English	Greek	Japanese
[s]	s	sa:⁵⁵ 'lightning'	*saw*	soba 'stove'	senaka 'back'
[k]	k	ku:³³ 'drum'	*goose*	kukla 'doll'	kodomo 'child'
[t]	t	ti:p³⁵ 'plates'	*danger*	tonos 'tuna'	teɡami 'letter'
[ts]	ts	tsi:t⁵⁵ 'tickle'	*	tsepi 'pocket'	tsukʲi 'moon'
[kʲ]	kj	kʲi:p³⁵ 'clip'	*gift*	kʲoskʲi 'kiosk'	kʲu:ri 'cucumber'
[ʃ]	S	*	*shop*	*	ʃika 'deer'
[tsʰ]	tsh	tsʰɛ: ⁵⁵ 'car'	*	*	*
[tʰ]	th	tʰi:n⁵⁵ 'sky'	*teeth*	*	*
[tʃ]	tS	*	*jet*	*	tʃu:ʃa 'injection'
[kʰ]	kh	kʰɔ:ŋ²¹ 'poor'	*cougar*	*	*
[ð]	D	*	*these*	ðakri 'tear'	*
[kʲʰ]	kjh	kʰi:u²¹ 'bridge'	*cube*	*	*
[dʒ]	dZ	*		*	dʒu: 'ten'
[d]	d	*	*	domata 'tomato'	denwa 'phone'
[ɡ]	g	*	*	ɡofreta 'candy bar'	ɡoma 'sesame'
[ç]	C	*	*	çoni 'snow'	çaku 'hundred'
[dz]	dz	*	*	dzami 'glass'	dzo: 'elephant'
[ɡʲ]	gj	*	*	ɡʲemɲa 'reins'	ɡʲu:nʲu: 'milk'
[θ]	T	*	*thin*	θ esi 'seat'	*
[kʷ]	kw	kʷɒ:⁵⁵ 'melon'	*(guano)*	*	*
[ɟ]	J	*	*	ɟiða 'goat'	*
[x]	x	*	*	xorta 'greens'	*
[kʷʰ]	kwh	kʷʰɒ:ŋ⁵⁵ 'frame'	*queen*	*	*
[tʃʰ]	tSh	*	*cheese*	*	*
[z]	Z	*	*zipper*	zimi 'dough'	*
[ɣ]	G	*	*	ɣonata 'knee'	*
[tʷʰ]	twh	*	*twin*	*	*
[tʷ]	tw	*	*(dwindle)*	*	*
[ʒ]	Z	*	*(jabot)*	*	*

We are interested in lingual obstruents because these are less transparent to the young language learner than are either labials (where there are clear visual clues to the place and manner of articulation) or glides (where the auditory feedback about lingual posture is available simultaneously with the somatosensory feedback). To master the articulation of the [t] in *tōfu* 'tofu', *tisshū* 'tissue', or *teɡami* 'letter', for example, a Japanese child must deduce not only that the initial silence corresponds to an interval when the tongue blade is pressed against the upper incisors to seal off the oral cavity, but also that the posture of the body of the tongue behind the seal is responsible for the variable second formant transitions after its release. To master the dorsal articulation of the initial in *kyōsō*, *kimono*, or *kemuri*, similarly,

the child must deduce that the more compact shape that differentiates this stop's burst from the [t] burst corresponds to a different lingual posture in which the tongue blade is not raised away from the tongue body, but instead is tucked down to allow the pre-dorsum to be tightly bunched up to contact the palate and seal off the oral cavity just behind the alveolar ridge.

Since this parsing of lingual gestures can be more or less difficult before different vowels, it is important to sample across a variety of contexts. Whenever possible, we elicited a target consonant in three words in each of five types of following vowel context, which we will call [i, e, u, o, a]. For Greek and Japanese, these types are just the five contrasting vowel phonemes (with "[u]"=[ɯ] in Japanese). For the other two languages, each of the context types included all phonemes in roughly the same region of the vowel space as the corresponding Greek or Japanese vowel. For example, the [e] category included both the long [ɛ:] of [tɛ:^{55}ti:^{21}] 'Daddy' and the short [e] of [tek^{55}si:^{35}] 'taxi' for Cantonese, and both the tense diphthongal [ej] of *cake* and the lax [ɛ] of *ketchup* for English.

Fig. 1 illustrates the resulting accuracy data. The dots in each panel show the transcribed accuracy of productions of all of the target dorsal stops for one of the languages, averaged by child and plotted as a function of the child's age. The overlaid lines show the predicted accuracy by age, as calculated by a logistic mixed effects regression model with age and dorsal stop type as fixed effects and child as the random grouping factor in an intercept-as-outcomes model (Raudenbush and Bryk 2002). The black solid lines graph the model coefficients for the "basic" velar stops and for each of the contrastively labialized and/or palatalized velar stops of the language, with the "basic" velar stops defined as the velars just in the back vowel environments where "basic" and "elaborated" articulations contrast in all of the languages. The dashed gray lines graph the model coefficients for the allophonically palatalized velars before front vowels.

As Fig. 1a shows, the Cantonese-acquiring children's productions were most accurate for [k] and least accurate for [k^{wh}], and the differences were larger for younger children. This pattern matches the patterns reported in other cross-sectional studies such as Cheung (1990) and So and Dodd (1995) even though those studies used different elicitation tasks and materials that sampled the consonants in only one or two words each. The apparent-time developmental pattern of different relative accuracies across the four age groups is congruent also with the real-time patterns reported for the ten children in Stokes and To's (2002) longitudinal study. The errors, too, are similar across studies, with misarticulated productions of [k^{w}] and [k^{h}] most typically transcribed as delabialization and deaspiration to [k], and misarticulated productions of [k^{wh}] most typically transcribed as simplifications in either or both of these dimensions of contrast.

The one pattern that is not simply a replication of earlier results is the lower accuracies for the allophonically palatalized stops relative to the "basic" velars before back vowels. This difference could not have emerged in earlier studies, which elicited dorsal stops in far fewer words than in the current experiment and so could not control for this conditioning by vowel context. Because the παιδολογος database includes productions of the Cantonese dorsal stops elicited in enough different words to be able to differentiate front vowel contexts from back vowel contexts, we can compare the relative accuracies in these two environments, to see that allophonically palatalized stops in Cantonese are relatively less accurate in the same way that they are in Greek and Japanese, the two languages with a fairly robust contrast between "basic" and palatalized velars before back vowels.
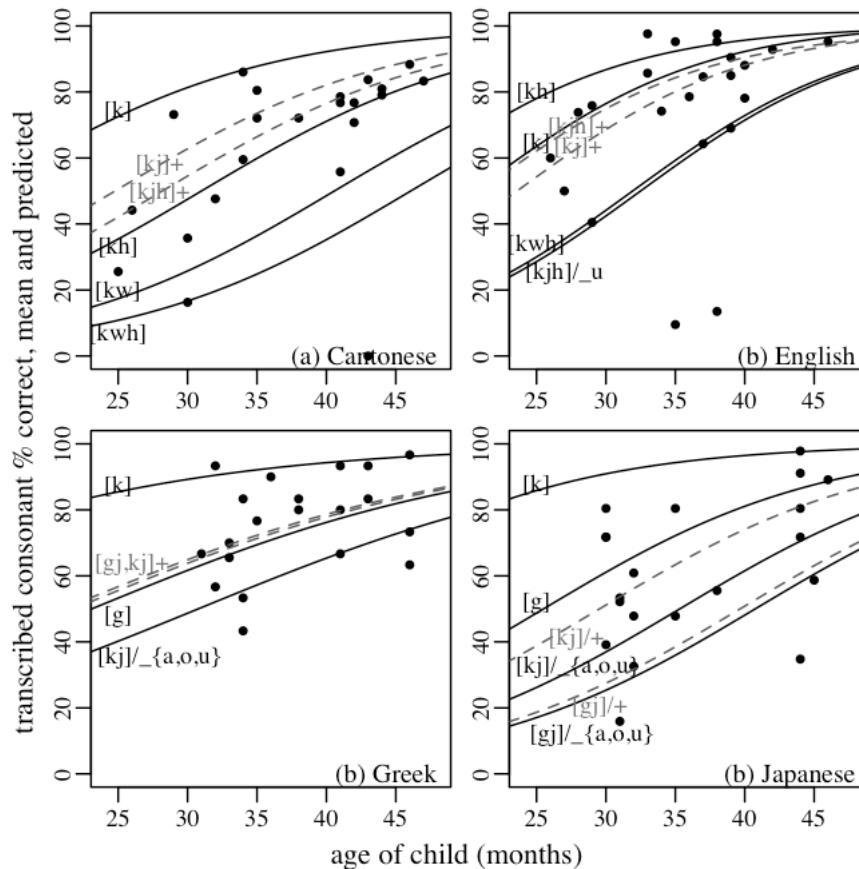

Figure 1. Mean transcribed accuracy of the dorsal stop types elicited in each language as a function of the child's age, with predicted accuracies for each stop type overlaid. (See Table 1 for the WorldBet symbol labels.)

As Fig. 1b shows, many of these trends also hold for the transcribed productions by the English-acquiring children. Accuracies were higher for the velar stops at the beginning of words such as *cougar* and *goose* (which are aspirated [kʰ] contrasting with a typically voiceless unaspirated [k] in the dialect of the children we recorded) in comparison to the labialized velar [kʷʰ]. Also, these children showed lower accuracy for the allophonically palatalized [kʲʰ] and [kʲ] in words such as *key* and *gift*, albeit not as low as for the contrastively palatalized [kʲʰ] in words such as *cupid*.

The one Cantonese trend that is not apparent in the analogous English stops in Fig. 1b is that there were very few de-aspiration errors, and when there are contrasting pairs of aspirated and unaspirated consonants, the aspirated stops had, if anything, somewhat higher accuracy rates rather than lower ones. Moreover, the errors in both cases were predominately errors of place of articulation ("velar fronting") rather than errors of phonation type. This is in keeping with results of both large-scale norming studies such as Smit et al. (1990) and of more focused longitudinal studies such as Macken and Barton (1980), the latter of which shows the English word-initial aspiration contrast to be mastered by 26 months in children with typical phonological development.

Finally, as Figs. 1c and 1d show, Greek- and Japanese-acquiring children were like the Cantonese-acquiring children in being most accurate for [k]. They were less accurate for voiced [ɡ] (showing devoicing as well as other errors) and also for allophonically palatalized [kʲ], and least accurate for contrastively palatalized [kʲ] and [ɡʲ]. While the patterns are similar between these two languages, it is somewhat difficult to relate them to results of previous research on either language. This is particularly true of Greek, where logopedic standards are only beginning to be established. However, the effect of palatalization is difficult to compare to earlier studies even for Japanese, since large-scale norming studies typically do not include contrastively palatalized [kʲ] or [ɡʲ] in the test materials and analyze the allophonically palatalized initial differently. For example, Nakanishi, Owada and Fujita (1972) phonemicize the voiceless dorsal before [i] as /kʲ/, but they include the allophonically palatalized dorsal before [e] as well the "basic" dorsal before [u, o, a] in their [k] category.[4] On the other hand, the error patterns in our study agree with theirs, as well as with what little information that we can find about errors for contrastively palatalized [kʲ] in more focused studies such as Tsurutani (2004). The most typical error for [kʲ] is the substitution of a postalveolar affricate.

3. Counting consonant types for cross-language comparisons

The difficulty we face in trying to compare the accuracy rates in Fig. 1d to those reported in Nakanishi et al. (1972) is a more general problem. In or-

der to compare accuracy rates across languages, we must specify which consonants types to compare. How should we analyze the plosive systems of Cantonese and Japanese in order to compare the relative accuracies in Fig. 1a and Fig. 1d? Should we equate Cantonese [kʰ] to Japanese [k] and Cantonese [k] to Japanese [ɡ], as suggested by cognate pairs such as Cantonese [kʰaː⁵⁵laːi⁵⁵ou⁵⁵kʰei⁵⁵] ~ Japanese *karaoke* and Cantonese [kiːt³³tʰaː⁵⁵] 'guitar' ~ Japanese *gitaa*? That would miss the generalizations that both groups of children make the fewest errors for voiceless unaspirated [k] and that phonation type errors are different, with Cantonese [kʰ] being deaspirated and Japanese [ɡ] being devoiced, results that are in keeping with the many other studies of the acquisition of aspiration and voicing contrasts reviewed in Kong (2009). Similarly, how should we analyze the plosive system of English in order to assess the relationship between the velar softening errors reported for younger Japanese children and the velar fronting errors that Ingram (1974), Stoel-Gammon (1996), and others note for English-acquiring children? Should we equate the front dorsal stop in English *key* and *cabbage* with the palatalized initial in the Japanese cognate forms [kʲiː] and [kʲabetsu]?

The list of word-initial obstruents in Table 1 shows the equivalence classes that we set up in order to be able to compare accuracy rates across languages. As shown in the table (as well as in the WorldBet IPA symbols that label the regression curves in Fig. 1), we have chosen to classify each stop as belonging to one of three phonation type categories — voiced stops (as in Greek [ɡofreta] and Japanese [ɡoma]), voiceless stops (as in Cantonese [kuː³³], English *goose*, Greek [kukla], and Japanese [kodomo]), and aspirated stops (as in Cantonese [kʰɔŋ²¹] and English *cougar*). Also, we have chosen to classify each dorsal stop as belonging to one of three place of constriction categories — the "basic" velar constriction type (as in Cantonese [kuː³³], English *goose*, and so on), a front or palatalized type (as in Cantonese [kʲiːp³⁵], English *gift* and *cupid*, Greek [kʲoskʲi] and [kʲiklos] 'circle', and Japanese [kʲuːri] and [kʲitsune] 'fox'), and a doubly articulated labialized velar type (as in Cantonese [kʷɒː⁵⁵] and English *queen*). We have chosen these three types in order to maximize the number of cross-language comparisons that can be made while minimizing the phonetic inaccuracies that are intrinsic to making any equation across languages (see note 1).

We have also been guided by the number of types that are implicit in the IPA chart, which are also used in most cross-language comparisons. For example, Lisker and Abramson (1964), Keating (1984), and many others have suggested that, while there are fine-grained differences across languages, there also seem to be recurring patterns in the distribution of voice onset time (VOT) values measured in word-initial stops which support a first rough split into just the three types that we have listed in Table 1. Moreover, this three-way classification into voiced stops (lead VOT),

voiceless stops (short lag VOT), and aspirated stops (long lag VOT) has been useful in understanding some of the observed cross-language commonalities in patterns of phonological acquisition. Specifically, the literature on first language acquisition of stop phonation types for the most part supports 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. The child thus generalizes this articulation independently of whether the particular prototype opposes the voiceless unaspirated stop to a voiced unaspirated stop (as in the Slavic and Romance languages), or to a voiceless aspirated stop (as in Danish). (Jakobson, 1941, p. 14)

As Macken (1980) points out, English seemed to be an exception to this generalization until it was recognized that the English "voicing" contrast is equivalent to the Danish opposition as described by Jakobson. In making this point, Macken could cite Lisker and Abramson (1964) and many other studies of North American (and Southern British) English dialects suggesting that, aside from Indian English, most standard varieties have aspirated stops (with long lag voice onset times) in contrast with unaspirated stops. Acoustic analyses of adult productions recorded in the second phase of the παιδολογος project (Kong, 2009) confirms that this characterization is true also of the Ohio dialect of the speakers that we recorded; all of the adult tokens of target consonants in words such as *teeth* and *cougar* had long lag VOT, as in the Cantonese aspirated stops, and only a small handful of tokens of target consonants in words such as *danger* and *gift* showed voicing lead. The Japanese voicing contrast, on the other hand, is closer to the Slavic or Romance opposition described by Jakobson. That is, our Japanese adult productions of target stops in words such as *tegami* and *kodomo* conform to the results of earlier studies such as Homma (1980) and Riney et al. (2007) in showing VOT values that are somewhat longer than those measured for the Cantonese unaspirated stops and English "voiced" stops but still considerably shorter than the VOT values in the Cantonese and English aspirated stops. Moreover, a large proportion of Japanese adult productions of target stops in words such as *denwa* and *goma* showed voicing lead, and children's productions of these stops with short lag VOT were often transcribed as having devoicing errors. Therefore, although the Japanese stop voicing contrast does not phonetically align exactly with the phonation type contrast in any of the other three languages, it seems less incorrect to equate it with the Greek voicing contrast than with the aspiration contrasts in the other two languages.

The reasoning that led to our three-way differentiation among a "basic" velar constriction, an "elaborated" front dorsal or palatalized dorsal constriction, and an "elaborated" labiovelar double constriction was similar. That is, evidence such as the mid-sagittal cineflourographic tracings in

Wada et al. (1970) might argue for a finer-grained differentiation for some of the languages. These tracings show that allophonic variation in place of constriction for Japanese dorsal stops in adult productions is not simply bimodal. Rather, the place of contact along the soft and hard palate varies continuously as a function of the following vowel, from the most posterior constriction before [o] to the most anterior constriction before [i]. This continuum is evident also in analyses of dorsal stop burst spectra reported in Arbisi-Kelm, Beckman, Edwards and Kong (2007), which show that Japanese adults' (and children's) bursts have peak frequencies ranging from lowest for [ko] to highest for [kʲi, kʲa, kʲo, kʲu], with different intermediate values for [ka], [ku], and [kʲe], in that order. However, we have no direct articulatory evidence about the dorsal constriction of the contrastively palatalized [kʲ] before back vowels. Moreover, such finer-grained variation is quite language-specific. The distribution of spectral peak frequencies for the Greek bursts examined by Arbisi-Kelm et al. (2007) suggests that the Greek dorsals have the most posterior constriction before [u] and much more front constrictions before [a] and [e] as compared to the analogous Japanese contexts. Further evidence of the very front constriction before [e] can be seen in the electropalatagraphic records in Nicolaidis (2001). At the same time, the burst spectra for English voiceless dorsal stops do not support any finer-grained specification than the binary differentiation between a back constriction in words such as *cougar* and *coat* and a front constriction in words such *cake*, *key*, and *cute*. Making just two groups for English is also in keeping with the review of extant X-ray data surveyed in Keating and Lahiri (1993). Therefore, in order to be able to make any comparisons at all for dorsal stops across the languages, we collapse the two types of palatalized stops that are differentiated in Figs. 1b-d to get just two dorsal types — the unmarked "basic" [k] versus a marked "elaborated" [kʲ].

We do not have space here to explain every equivalence class that we chose to set up. Our current choices are necessarily based on work in progress, such as Li (2008) for voiceless sibilant contrasts, and Schellinger (2008) for stridency contrasts. Eventually, all of the classes (like the classes for phonation type and for dorsal constriction type just discussed) should be based not just on extensive reading in the literature on these languages, but also on acoustic analyses of the adult and child productions in the παιδολογος database, supplemented by cross-language speech-perception experiments. In the meantime, the medium-grained differentiations that we have set up seem adequate for our current purpose, which is to try to understand what cross-language similarities and differences in accuracy rates, such as the rates plotted in Fig. 1, can tell us about paths of convergence for children who are learning the words of these four different languages. That is, setting up these equivalence classes allow us to compare accuracies and also to compare the type frequencies of the consonants in word-initial posi-

tion, in order to assess the lexical support for each consonant in a comparable way across languages.

Once we had set up the equivalence classes, we measured each consonant's frequency in each language by counting the number of words beginning with that consonant in a reasonably large wordlist and dividing by the total number of words. The Cantonese list is the 33,000 words segmented by rule from the Cantonese-language portion of a transliterated corpus of Chinese newspapers (Chan and Tang, 1999). The English list is the 19,321 forms counted after collapsing homophones in the intersection of several pocket dictionaries (Pisoni et al., 1985). The Greek list is the 18,853 most frequent forms in a large morphologically tagged corpus of newspaper texts (Gavrilidou et al., 1999). The Japanese list is the subset of 78,801 words in the NTT database for which familiarity ratings are available (Amano & Kondo, 1999).[5] The list of word-initial obstruents in Table 1 is ordered by average frequency in these four lists.[6] The order in the table is congruent with what we expect from studies such as Lindblom and Maddieson (1988); the consonants at the top include the most unmarked "basic" types [s], [k], and [t], while those further down include "elaborated" types such as [z] and [θ], and "complex" types such [k^{wh}]. This congruence suggests that our measure of average type frequency could be adapted to be a proxy measure of rank on a language-general markedness scale in a more detailed quantitative assessment of the influence of markedness on relative consonant accuracy in each language. Before developing such a measure, however, we will first assess the extent to which accuracy patterns are similar across pairs of languages, and whether similarities can be accounted for by corresponding similarities in the language-specific consonant frequencies.

4. Correlating consonant patterns across pairs of languages

The link between these pairwise correlational analyses and the larger question is the following chain of arguments. First, if language-general markedness influences the acquisition of lingual obstruents, the accuracy rates for consonants shared by any two languages should be correlated. Second, if this influence is direct, then the correlation between consonant accuracy rates should be independent of any correlation between their frequencies in the two lexicons. However, if the influence is only indirect, because it is mediated by the influence of markedness on the lexicon through sound change, then the frequencies of these shared consonants in the two languages' lexicons should also be correlated, with a shared variance that is at least as large as the shared variance for the accuracy rates. Also, the order of the shared consonants when ranked by the correlated frequencies should match the order of shared consonants when ranked by their accuracy rates.

Fig. 2 shows the six sets of correlations that we did, with plots for the accuracy relationships on the left and plots for the frequency relationships on the right. Each number along the x- or y-axis in the plots on the right is the natural logarithm of the count of words beginning with the consonant divided by the total number of words in the list for that language. We use the log ratio rather than the raw ratio, because this effectively weights a percentage change at the low-frequency end of the distribution more heavily than the same percentage change at the high-frequency end, reflecting the intuition that changes at low frequencies are more consequential than changes at high frequencies. For example, if nine words that are borrowed into a language's lexicon in a situation of language contact raise the frequency of a particular consonant from 1 to 10, this is a much more substantial change to the system than if the frequency were raised from 201 to 210.

The number in the lower left corner of each plot is the r^2 value returned by a Pearson product-moment correlation. The number is in black if the correlation was significantly different from no correlation at $\alpha=0.05$, and it is in gray (and surrounded by parentheses) if the correlation was not significant. The r^2 value is equal to the R^2 in a regression model, and indicates the proportion of variance that is shared between the correlated variables.

The language pairs are ordered in the rows of Fig. 2 by the r^2 for the accuracy rates, which ranges from a high of 0.82 (for the comparison between the Cantonese- and Greek-learning children) to a low of 0.02 (for the comparison between the English- and Japanese-learning children). In each of these six plots, the "basic" consonants [k] and [t] are in the upper-right quadrant of the graph (indicating high accuracy in both groups of children) and there is a tendency for the least accurate consonants in both languages to be "elaborated" or "complex" obstruents such as [ts], [ð], or [k^{wh}]. Also, each of the first four of these accuracy correlations is significantly better than no correlation at the 0.05 level.

By contrast, only one of the six correlations between consonant frequencies is significant at $\alpha=0.05$. This is the correlation between the relative frequencies in the Greek and Japanese wordlists for the ten lingual obstruent types shared between these two languages. However, the rank orders for the shared consonants in the two plots in this second row of the figure are not identical. Most notably, [s] is among the most frequent of lingual obstruents in both lexicons and [ɡ^j] among the least frequent, but [s] is not more accurate than [ɡ^j] in the productions of either group of children.
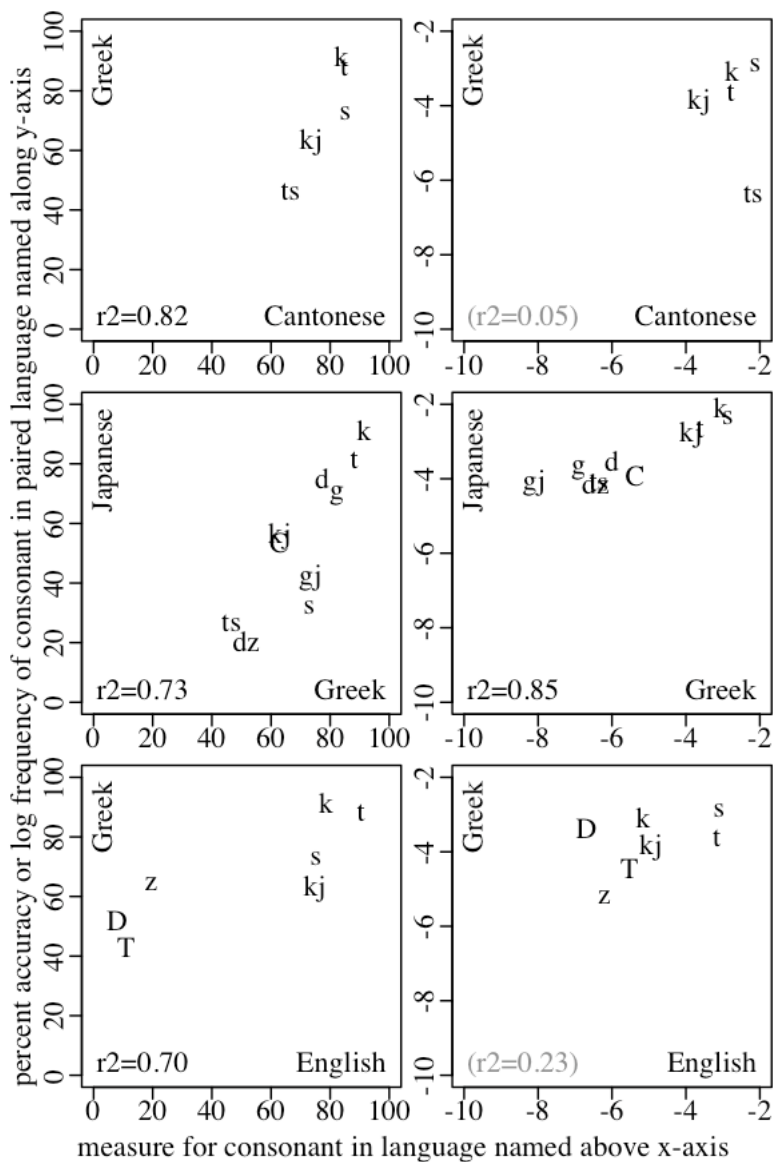
Figure 2. Relationships between the transcribed accuracy rates (left) and the log frequencies in the lexicon (right) for word-initial pre-vocalic consonant types shared by the each language pair, with correlation coefficients, plotted in black when the relationship is significant at the *p*<0.05 level and otherwise in gray. Plotting characters are the WorldBet (Hierony-mous, 1993) symbols used by the transcribers (see Table 1).
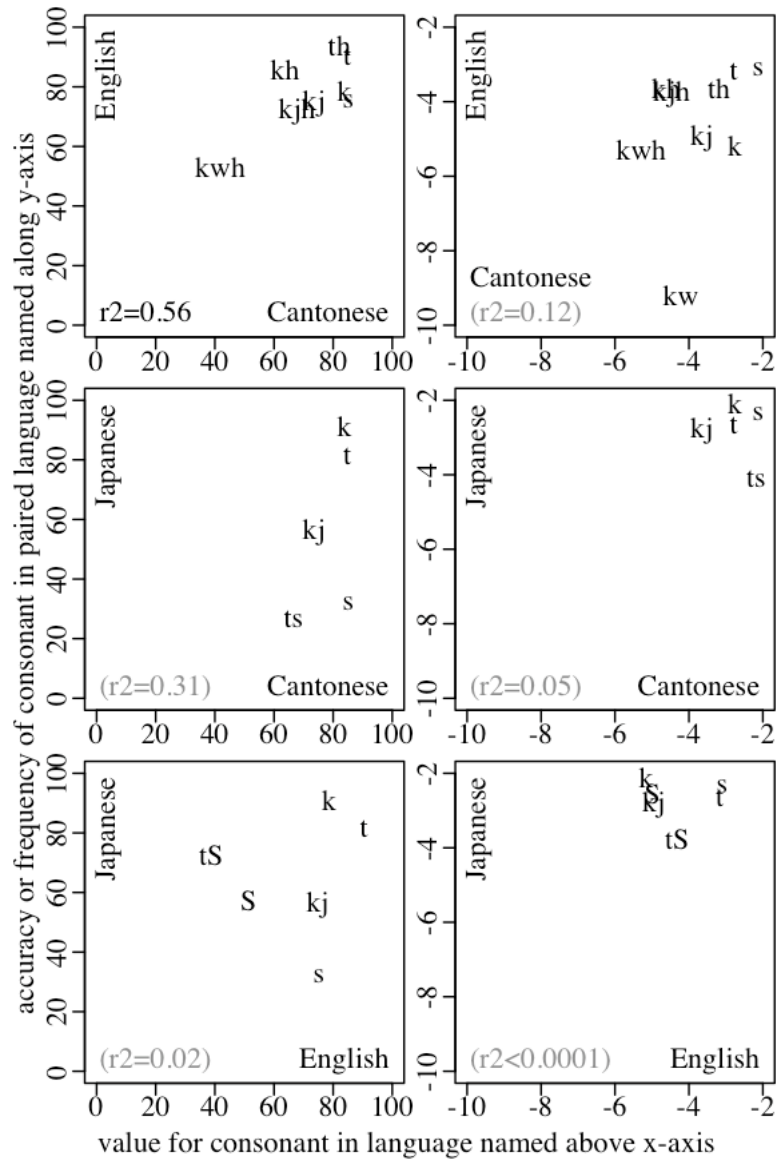
Figure 2. cont.

Our use of correlational analyses here is similar to the strategy adopted by Lindblom and Maddieson (1988) to explore phonetic universals for consonant systems. They correlate frequency of attestation in the inventories of the world's languages with the sizes of the consonant inventories in which different consonant types are attested, to find evidence for grouping consonant types into sets along a markedness scale from "basic" types such as [t, k], through types with "elaborated" articulations such as [k^h, k^w, k^j], to

"complex" types that combine two or more elaborated articulations such as [kʷʰ, kʲʰ]. They summarize their results saying: "Initially system 'growth' occurs principally in terms of basic consonants. Once these consonant types reach saturation, further growth is then achieved first by adding only elaborated articulations, then by invoking also complex segment types." (p. 70) This summary could describe the relative accuracy rates for the consonant types shown in Fig. 1. The smallest inventories of the youngest children include a fairly robust [k]. The larger inventories of slightly older children incorporate the elaborated types [kʲ] or [kʷ] and only the oldest have the complex type [kʷʰ]. This tendency for "basic" types to be produced accurately before "elaborated" types is also evident in the overall accuracy rates in Fig. 2. Correlations in accuracy rates between any two groups of children, then, suggest the influence of a universal scale of phonetic markedness. The general lack of correspondence to the arrangement of frequencies supports an interpretation of the influence as being, to some extent, direct.

The pair-wise correlational analyses in Fig. 2 have the advantage of being symmetrical — i.e., neither language is designated as contributing the independent variable and hence implicitly as being more representative of universal tendencies. However, these analyses cannot include marked sounds that are not shared. For example, the Cantonese and English aspirated stops cannot be included in pair-wise comparisons with Greek or Japanese, and the Greek voiced dorsal fricatives [ʝ] and [ɣ] cannot be included in any pair-wise comparison. Thus, these analyses may underestimate the role of markedness in shaping the relative frequencies of different consonant types in the course of language change. Our second set of analyses was intended to address this limitation.

5. Regression analyses using mean frequencies

For these analyses, we derived a measure of relative markedness for each consonant in a language's inventory by averaging the consonant's frequencies across the other languages, using 0 as the frequency for any language in which a consonant is not attested. This measure is like the average frequencies that we used to order the consonants in Table 1 except that it is customized to each target language by excluding the frequency for that language. We used this "other-language mean frequency" in the twelve regression models shown in the panels of Fig. 3, with the R^2 value in the lower-right corner given in black if a model accounts for a significant proportion of the variance in the y-axis variable and in gray otherwise.
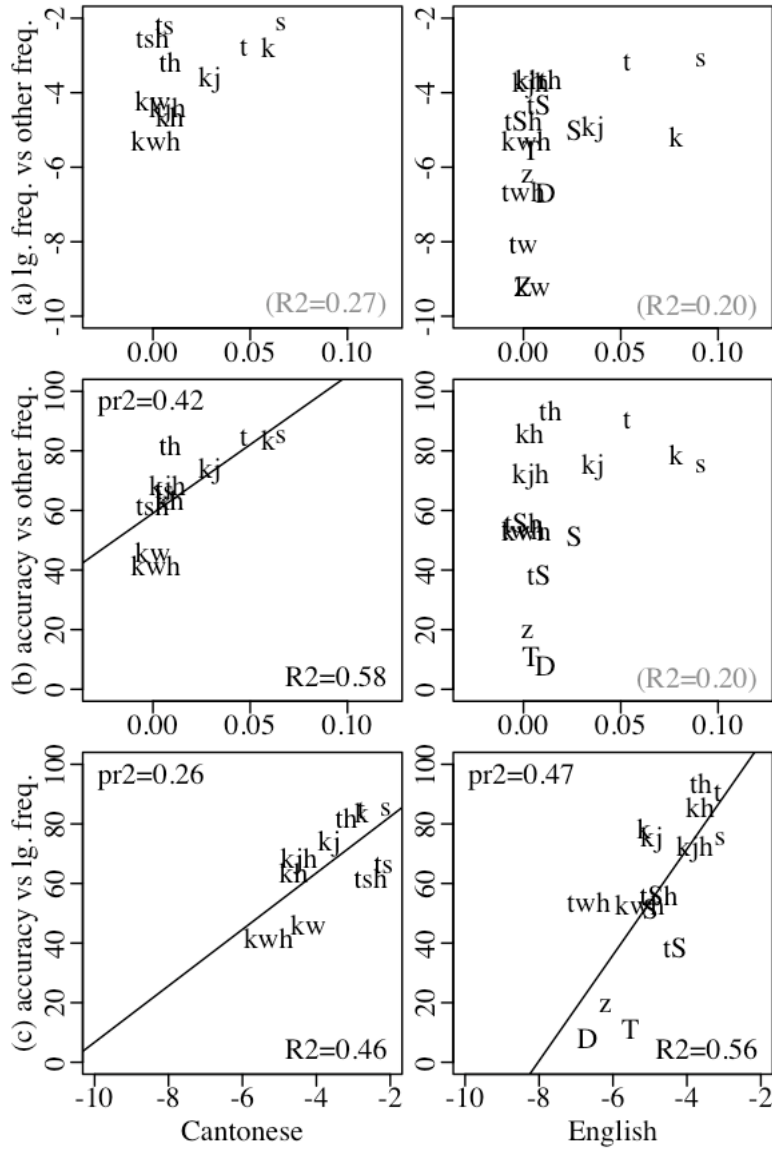
Figure 3. Scatterplots for (a) language-specific log frequency and (b) children's production accuracy rates as functions of mean frequency in the other three languages, and for (c) accuracy as a function of language-specific log frequency. Regression curves are drawn and R^2 values are shown in black for regression models that are significant at the 0.05 level.
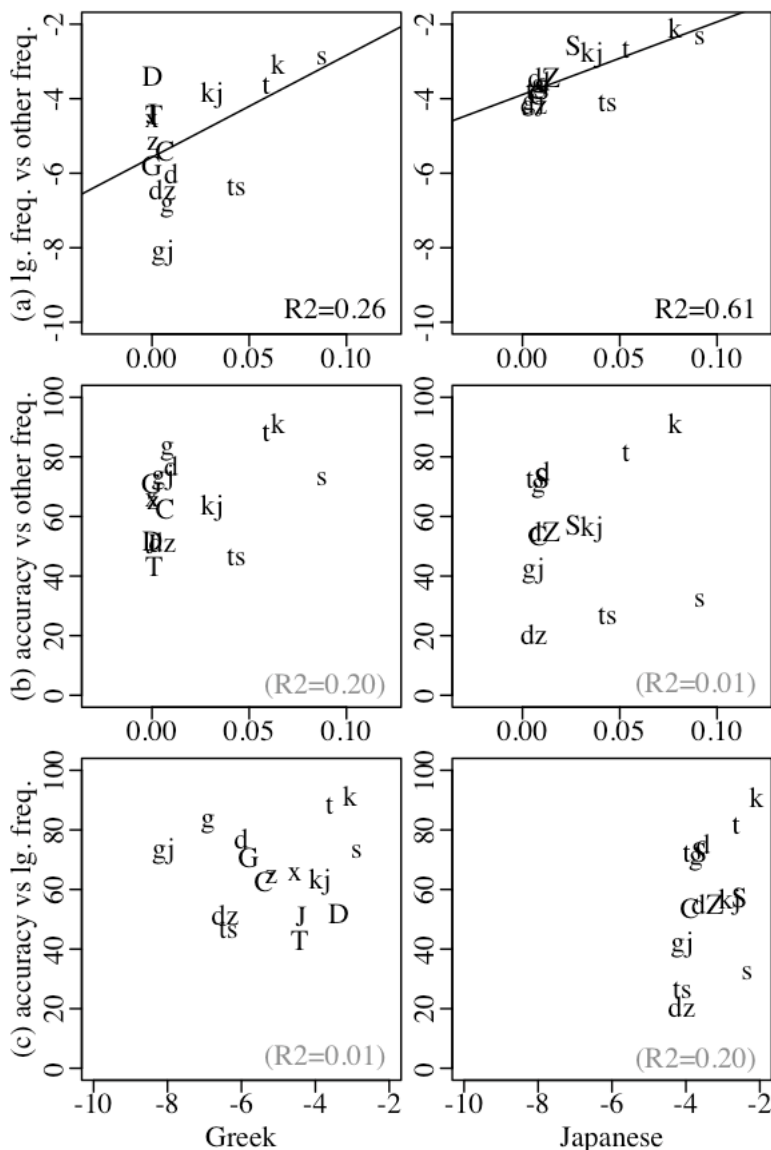
Figure 3. cont.

First, we regressed the log frequency in each language's lexicon against the mean frequency in the other three languages. This is the (a) panel in each column of Fig. 3. These regression models estimate the influence of language-general markedness on the lexical support for each lingual obstruent in the target language's consonant inventory. Mean other-language frequency accounted for at least 20% of the variance in log frequency for each language, and the relationship was significant in Greek and Japanese.

Second, we regressed the accuracy rate in the productions by the children acquiring that language against the mean frequency in the other languages. This is the (b) panel in each column of Fig. 3. These regression models estimate the influence of language-general markedness on the accuracy of young children's productions of the target language's obstruents. There was a significant relationship only for Cantonese. The pr^2 value shown in the upper-left corner of the panel for Cantonese is the amount of variance in accuracy accounted for by the measure of language-general markedness after the relationship in panel (a) between language-general markedness and language-specific frequency is partialled out.

Finally, we regressed the accuracy rate in the productions by the children acquiring each language against the log frequency in that language's lexicon. This is the (c) panel in each column of Fig. 3. These regression models estimate the effect of language-specific lexical support on children's production accuracies. There was a significant relationship in Cantonese and in English. The pr^2 value shown in the upper-left corner of the panels for each of these two languages is the amount of variance in the accuracy rates accounted for after the relationship in panel (a) between the language-specific log frequency and the other-language mean frequency is partialled out.

The R^2 values in the (a) panels of Fig. 3 suggest that, despite the negative evidence in the pair-wise correlations for language-specific frequency in Fig. 3, there is some support for the idea that an unmarked consonant tends to occur in many words and that a marked consonant that is attested in few languages tends to occur in fewer words in a language that has it in its inventory. However, what is perhaps more illuminating in these regressions is that outliers can often be interpreted in terms of sound changes that might affect lexical frequencies in ways that are opposite to the predicted relationship between markedness and frequency. For example, the aspirated stops of English are more frequent than one would expect for this "elaborated" phonation type, but Grimm's Law and its aftermath helps us understand their exceptional status. Similarly, the voiced fricatives of Greek (especially [ð]) are more frequent than we might expect unless we remember that they are reflexes of Proto-Indo-European *b, *d, *g^j, *g. The [ts] of Cantonese also is very frequent, because of a historic merger that neutralized contrasts in sibilant place of articulation. By contrast, [ts] is rare in Japanese because it emerged out of a diachronic change similar to the one that produced the affricate allophones of [t] and [d] that occur before [i, y, u] in Quebec French. A phonotactic restriction that also holds for [s] before [i] results in the contemporary pattern of [ts] occurring almost exclusively in the context of a following [u]. (This is the opposite distribution to the one in Cantonese, where the high back vowel is the only following context

where [ts] does *not* occur, because of a sound change fronting [u] to [y] in the context of a preceding coronal consonant.)

Modern Greek has no such phonotactic restriction, but Greek [ts] also is lower in frequency than Cantonese [ts] because of the different historical sources for affricates in the two languages. Here we invoke Blevins's (2006) distinction between "natural" system-internal sources of universal patterns in sound change and "unnatural" or system-external sources of systemic change. In this typology, Greek [ts] is not the result of regular system-internal change. Rather, Joseph and Philippaki-Warburton (1987) describe the two Modern Greek affricates as gradually emerging in the 10th or 12th century via borrowings from other languages (e.g., [tsepi] 'pocket' < Turkish *cep*) and from dialects with velar softening. In Cantonese, by contrast, some words beginning with [ts] also are loanwords (e.g., [tsɛ:m^{55}] 'jam'), but the majority are not. That is, as far back as we can reconstruct the Chinese consonant system, there seem to have been affricates at both the dental/alveolar place and at one or two post-alveolar places of articulation. The single place of articulation in modern Hong Kong Cantonese reflects a series of mergers among these places, as illustrated by correspondences in Cantonese and Mandarin pairs such as [tsɐu^{35}]~[tɕo^{214}] 'wine' and [tsɐm^{55}]~[tʂɤn^{55}] 'needle'. These comparisons highlight for us the many different social factors that can complicate the relationship between markedness in language acquisition and markedness in language change.

If our interpretation of such outliers in the regressions between language-specific log frequency and other-language mean frequency is correct, we can use relationships such as those plotted in the (a) panels of Fig. 3 to begin to disentangle the different paths of convergence in children's accuracy relationships. That is, let us assume that, to a first approximation, frequency relationships will be more similar across languages when they result from common across-the-board sound changes. And they will be more particular to a given language when they result primarily from borrowings and more localized analogical changes within a language. If this assumption is correct, then when we regress accuracy against the mean other-language frequency, we get a liberal estimate of the influence of universal markedness on consonant acquisition, and we can partial out the correlation between mean other-language frequency and language-specific frequency to get a more conservative estimate. Also, when we regress accuracy against language-specific log frequency, we get a liberal estimate of the role of historical contingencies in the cultural evolution of the lexicon, and we can partial out the correlation between language-specific frequency and mean other-language frequency to get a more conservative estimate.

The (b) and (c) panels of Fig. 3, then, are these estimates. As can be seen from the numbers in these eight panels, the predictive value of the models differs across languages, with Greek and Japanese showing little or

no variance accounted for by either regression, and Cantonese and English showing at least half of the variance in accuracy predicted by one of the frequency measures. It is interesting, too, that the latter two languages differ in terms of which regression model is more predictive of the accuracy relationships among the consonants that children must master in order to become fluent adult speakers of the language. For Cantonese-learning children, the mean other-language frequency was more predictive. The panel (b) regression accounted for 42% of the variance even after the relationship between the mean other-language frequency and the Cantonese-specific frequency is partialled out, suggesting a strong direct influence of universal markedness on Cantonese consonant acquisition. By contrast, for English-learning children, the language-specific frequency in panel (c) was far more predictive, and other-language frequency in panel (b) accounted for no more of the accuracy for English-learning children than it did for the Greek-learning children.

Also, as with the patterns in the (a) panels of Fig. 3, we again find interesting outliers that seem to be related to historical contingencies that have conspired to make for some unexpected patterns of lexical frequencies. In the (c) panel for Greek, for example, [ð], [j̇], and [ɣ] again stand out. These voiced fricatives, which are rare cross-linguistically, developed historically in Greek from voiced stops, and are much less accurate than predicted from their high frequencies relative to the voiced stops that occur in loanwords such as [gol] 'goal'.

Also, in both Cantonese and English we can compare the accuracy of each aspirated stop to the accuracy predicted by the regressions in the lower two panels for these languages. While aspirated stops are not uncommon across languages, they are less common than voiceless unaspirated stops, which seem to be attested in every spoken language. This relationship is mirrored in the Cantonese lexicon. That is, in Cantonese, each aspirated stop is attested in fewer words than its unaspirated counterpart, because of the relative frequencies of the four Middle Chinese tones and how these tones conditioned the outcome of the tone split that merged the Middle Chinese voiced stops (still attested in the Wu dialects) together with either the voiceless unaspirated stops or together with the aspirated stops. The relative accuracies also conform to the predicted relationship, in both the (b) and the (c) panel regressions. In English, by contrast, the aspirated stops are more frequent than their unaspirated counterparts, having developed historically from more "basic" voiceless unaspirated stops, which were frequent not just in the native Germanic vocabulary but also are represented in a large number of loanwords from French. The English aspirated stops, then, are far more accurate than predicted by their low mean frequencies in the other languages in panel (b), but conform nicely to the relationship between accuracy and language-specific frequencies in panel (c).

6. Conclusions

In summary, the relationship between convergent patterns in acquisition noted in work such as Locke (1983) and the markedness hierarchies uncovered in work such Lindblom and Maddieson (1988) and Guion (1998) is very complex. Each child acquires the phonology of the specific ambient language, and even the most similar consonants across languages will be at best close analogues rather than homologues. So generalizations across different lexicons can lead to variation in development across languages. At the same time, the huge variation in the R^2 values for the relationships between accuracy and language-specific lexical frequency in the (c) panels of Fig. 3 suggests that other factors besides frequency also contribute to the differences in the relative order of mastery of obstruents across languages. Order of mastery cannot be explained entirely by the relative frequencies of the different consonants in the lexicon that the child is trying to grow, whether or not those frequencies come from natural sound changes. Neither can it be explained completely in terms of the pan-species biological and neural factors that constrain what consonants the child can produce and perceive reliably at different stages of maturation. Thus it would be a mistake to posit a deterministic relationship between any universal markedness hierarchy and order of mastery in speech acquisition. The results shown in Figs. 2 and 3 must be inconclusive, because the correlations and regressions are tests of predictions of very simple deterministic models.

Thus, the exercise of looking at these overly simple numerical tests was useful primarily because it clarifies the limits on the kinds of questions about markedness that can be addressed by looking at speech development only in one group of children, such as the group of typically-developing children who are acquiring English. All children are affected by the same developmental universals arising from the biological endowment of the species, but part of that endowment is that children have "powerful skills and motivations for cooperative action and communication and other forms of shared intentionality" such that "regular participation in cooperative, cultural interactions during ontogeny leads children to construct uniquely powerful forms of cognitive representation" (Moll and Tomasello 2007). That is, human babies differ from infants of other primate species in their capacity to grasp the phonetic intentions of the speakers in their environments and to develop the robust, language-specific articulatory-motor, auditory-visual, and higher-order structural representations that they need to develop in order to be inducted into a particular language community. Many developmental differences across children acquiring different languages arise from the fact that mastery of a particular system of consonants means mastery of the lexicon that is shared by the culture of speakers who share the language. Because this lexicon is a unique product of the lan-

guage's history, children who are acquiring different languages will build different cognitive representations for "the same consonants" because this "sameness" is an identity by analogy, not by homology.

Correspondence e-mail address: mbeckman@ling.ohio-state.edu

Notes

1   We use IPA symbols set off in square brackets as a shorthand way to refer to the consonant sounds and vowel contexts that we compare across the four languages examined in this paper. Our decision to use square brackets rather than slashes (or some other typographical device such as boldface type, as in Ladefoged 2001) is essentially aesthetic. We do not intend to imply that sounds that are transcribed in the same way in any two languages have exactly the same phonetic value. Indeed, our examination of acoustic patterns and perceptual responses across these four languages lead us to agree with Pierrehumbert, Beckman and Ladd (2000: 286) when they say, "there is no symbolic representation of sound structure whose elements can be equated across languages" because "phonological inventories only exhibit strong analogies."

2   We follow Lindblom and Maddieson (1988: 70-71) and others in using the terms "marked" versus "unmarked" as convenient cover terms for a complex of scalar properties that have been implicitly assumed or explicitly invoked in surveys of commonly attested sound changes and synchronic sound patterns across languages by a long line of researchers (e.g., Grammont 1933; Trubetskoî 1958; Locke 1983). The terms "basic" versus "complex" and "elaborated" also are from Lindblom and Maddieson (1988), and allude to their to interpretation of the marked/unmarked continuum as a "scale of increasing *articulatory complexity*" which they set up in order to evaluate the numerical relationship between consonant or vowel inventory size and the relative likelihoods of particular segments to in an inventory. (As they put it, "Small paradigms tend to exhibit 'unmarked' phonetics whereas large systems have 'marked' phonetics.") However, they also point to differences between vowel inventory effects and consonant inventory effects, saying that, "Consonant inventories tend to evolve so as to achieve maximal perceptual distinctiveness at minimum articulatory cost." See Macken and Ferguson (1981) and Beckman, Yoneyama, and Edwards (2003), among others, for more extensive discussion of the various types of properties that have been invoked in the acquisition literature.

3   We also elicited [kʰa] or [ka] and two other CV sequences in three sets of nonsense words, using a picture of a "nonsense" item for that culture. We included nonwords in order to test the feasibility of using picture-prompted word repetition in an ecologically more natural nonword repetition task, in preparation for the second phase of the project, where we make a more focused comparison of selected target sequences that have different phonotactic probabilities across the four languages. However, since there were only a handful of nonwords in these first phase recordings, we exclude them in calculating the accuracy rates here.

4   This grouping reflects the phonemicization implicit in the Japanese *hiragana* and *katakana* syllabaries, which write [kʲa], [kʲu], and [kʲo] with the symbol for [kʲi] followed by the (subscript) symbol for [ja], [ju], or [jo], respectively. This orthographic pattern is the basis for the very abstract phonemic analysis assumed in the UPSID list of consonants for Japanese, which includes none of the initial obstruents in words such as [tʃuːʃa] 'injection', [çaku] 'hundred', and [kʲoːso] 'race' but instead analyzes these as allophones (before an "underlying" [j]) of [t], [h], and [k], respectively. The UPSID analysis of the initial consonants in Greek [çoni] 'snow', [kʲoskʲi] 'kiosk', and so on, also are attributed to an "underlying" medial [j] or [i]. Arguments for the underlying [j] or [i] refer to morphophonological alternations in stem-final position and to adult speakers' strong metalinguistic awareness of the abstract (semi)vowel, which is represented in the orthographic form of both the stem-final consonants and the non-alternating root-initial consonants. However, two- and three-year-old children are not typically literate and they typically do not have the adult-sized lexicons that might support a restructuring of root-initial phonetic cohort groupings by analogy to stem-final morphophonological groupings. We therefore choose to use more narrow transcription classes, so that each of these two languages is analyzed as having a larger consonant inventory than in the UPSID listing.

5   Although our current frequency analyses are based on these adult wordlists, we have also begun to analyze frequencies in wordlists extracted from transcriptions of a corpus of child-directed speech that we recorded from mothers (or other primary caretakers) of 10 one-year-old children in the same dialect areas as the children that we recorded for each of the languages. Analyses completed to date show strong correlations between the relative frequencies in the two types of wordlists, with the few outliers at the low-frequency end of the scale.

6   As noted, the wordlists that we used are of varying size, ranging from less than 20,000 words for English to more than 70,000 words for Japanese. While this suggests that Japanese frequencies should dominate in determining the order, [k] rather than [s] is the most frequent consonant in Japanese. Also, in another analysis, reported in a paper presented at the 11th Conference on Laboratory Phonology, we ranked the consonants by frequency within each language, and then calculated the mean of these ranks. The mean frequencies used here are strongly correlated with the mean ranks used earlier ($R^2 = 0.84$, $p < 0.001$).

Acknowledgement

References

Amano, Shigeaki & Tadahisa Kondo. 1999. *Lexical properties of Japanese*. Tokyo, Japan: Sanseido.

Arbisi-Kelm, Timothy, Mary E. Beckman, Jan Edwards & Eunjong Kong. 2007. Acquisition of stop burst cues in English, Greek, and Japanese. Poster presented at the Symposium on Research in Child Language Disorders, Madison, WI, June 7–9, 2007.

Beckman, Mary E., Benjamin Munson & Jan Edwards. 2007. Vocabulary growth and the developmental expansion of types of phonological knowledge. In Jennifer Cole & José Ignacio Hualde (eds.), *Laboratory Phonology 9*, 241-264. Berlin: Mouton de Gruyter.

Beckman, Mary E., Kiyoko Yoneyama & Jan Edwards. 2003. Language-specific and language-universal aspects of lingual obstruent productions in Japanese-acquiring children. *Journal of the Phonetic Society of Japan* 7. 18–28.

Blevins, Judith. 2006. A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32(2). 117–166.

Chan, Shui Duen & Zhixiang Tang. 1999. A quantitative analysis of lexical distribution in different Chinese communities in the 1990's. *Yuyan Wenzi Yingyong [Applied Linguistics]* 3. 10–18.

Cheung, Pamela. 1990. The acquisition of Cantonese phonology in Hong Kong — a cross-sectional study. B.Sc. thesis, University College London.

Edwards, Jan & Mary E. Beckman. 2008a. Some cross-linguistic evidence for modulation of implicational universals by language-specific frequency effects in phonological development. *Language Learning and Development* 4 (1). 122–156.

Edwards, Jan & Mary E. Beckman. 2008b. Methodological questions in studying consonant acquisition. *Clinical Linguistics and Phonetics* 22(12). 937–956.

Ferguson, Charles A. & Carol B. Farwell. 1975. Words and sounds in early language acquisition. *Language* 51(2). 419–439.

Gavrilidou, Maria, Penny Labropoulou, Elena Mantzari & Sophia Roussou. 1999. Specifications for a computational morphological lexicon of Modern Greek. In Amalia Mozer (ed.), *Greek Linguistics 97, Proceedings of the 3rd International Conference on the Greek Language*, 929–936. Athens: Ellinika Grammata.

Gould, Stephen Jay. 1989. *Wonderful Life: The Burgess Shale and the Nature of History*. New York: W. W. Norton & Co.

Grammont, Maurice. 1933. *Traité de phonétique*. Paris: Delagrave

Guion, Susan Guignard. 1998. The role of perception in the sound change of palatalization. *Phonetica* 55(1). 18–52.

Hieronymus, James L. 1993. ASCII phonetic symbols for the world's languages: WorldBet. Technical Report, AT&T Bell Laboratories.

Homma, Yayoi. 1980. Voice onset time in Japanese stops. *Onsei Gakkai Kaiho* 163. 7–9.

Ingram, David. 1974. Fronting in Child Phonology. *Journal of Child Language* 1. 233–241.

Jakobson, Roman. 1941. *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala: Almqvist & Wiksell. (Page numbers for quotations are from the 1968 translation *Child language, aphasia and phonological universals*. The Hague: Mouton.)

Joseph, Brian D. & Irene Philippaki-Warburton. 1987. *Modern Greek*. London: Croom Helm Publishers.

Keating, Patricia A. 1984. Phonetic and Phonological Representation of Stop Consonant Voicing. *Language* 60(2). 286–319.

Keating, Patricia A. & Aditi Lahiri. 1993. Fronted velars, palatalized velars, and palatals. *Phonetica* 50. 73–101.

Kirby, Simon, Mike Dowman & Thomas L. Griffiths. 2007. Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences of the U.S.A.* 104(12). 5241–5245

Kong, Eunjong. 2009. The development of phonation-type contrasts in plosives: Cross-linguistic perspectives. Ph.D. diss., Department of Linguistics, Ohio State University.

Ladefoged, Peter. 2001. *Vowels and consonants: An introduction to the sounds of languages*. Oxford: Blackwell.

Li, Fangfang. 2008. The phonetic development of voiceless sibilant fricatives in English, Japanese, and Mandarin. Ph.D. diss., Department of Linguistics, Ohio State University.

Lindblom, Björn & Ian Maddieson. 1988. Phonetic universals in consonant systems. In Larry M. Hyman & Charles N. Li (eds.), *Language, speech, and mind: Studies in honour of Victoria A. Fromkin*, 62–78. New York: Routledge.

Lisker, Leigh & Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20. 384–422.

Locke, John L. 1983. *Phonological acquisition and change*. New York: Academic Press.

Macken, Marlys A. 1980. Aspects of the acquisition of stop systems: A cross-linguistic perspective. In Grace H. Yeni-Komshian, James F. Kavanagh, & Charles A. Ferguson (eds.), *Child phonology*, *Vol. I Production*, 143–168. New York: Academic Press.

Macken, Marlys A. & David Barton. 1980. The acquisition of the voicing contrast in English: A study of Voice Onset Time in word-initial stop consonants. *Journal of Child Language* 7(1). 41–74.

Macken, Marlys A. & Charles A. Ferguson. 1981. Phonological universals in language acquisition. *Annals of the New York Academy of Sciences* 379. 110–129.

Maddieson, Ian & Kristin Precoda. 1989. Updating UPSID. *Journal of the Acoustical Society of America* 86. S19. [Description of the UPSID-PC DOS-based software at: http://www.linguistics.ucla.edu/faciliti/sales/software.htm]

Moll, Henrike & Michael Tomasello. 2007. Cooperation and human cognition: The Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1480). 639–648.

Nakanishi, Yasuko, Kenjiro Owada & Noriko Fujita. 1972. Kōon kensa to sono kekka ni kansuru kōsatsu [Results and interpretation of articulation tests for

children]. *RIEEC Report* [*Annual Report of the Research Institute for the Education of Exceptional Children, Tokyo Gakugei University*] 1. 1–41.

Nicolaidis, Katerina. 2001. An electropalatographic study of Greek spontaneous speech. *Journal of the International Phonetic Association* 31(1). 67–85.

Ohala, John J. 1992. What's cognitive, what's not, in sound change. In Günter Kellerman & Michael D Morrissey (eds.), *Diachrony within synchrony — language history and cognition: Papers from the international symposium at the University of Duisburg, 26-28 March 1990*, 309–355. Frankfurt: Peter Lang.

Pierrehumbert, Janet, Mary E. Beckman & D. Robert Ladd. 2000. Conceptual foundations of phonology as a laboratory science. In Noel Burton-Roberts, Philip Carr & Gerard Docherty (eds.), *Phonological knowledge: Its nature and status*, 273–303. Oxford: Oxford University Press.

Pisoni, David B., Howard C. Nusbaum, Paul Luce & Louisa M. Slowiacek. 1985. Speech perception, word recognition, and the structure of the lexicon. *Speech Communication* 4. 75–95.

Pye, Clifton, David Ingram & Helen List. 1987. A comparison of initial consonant acquisition in English and Quiché. In Keith E. Nelson & Ann Van Kleeck (eds.), *Children's Language, 6,* 175–190. Hillsdale, NJ: Lawrence Erlbaum.

Raudenbush, Stephen W. & Anthony S. Bryk. 2002. *Hierarchical linear models: Applications and data analysis methods*, 2nd edition. Thousand Oaks, CA: Sage Publications.

Riney, Timothy James, Naoyuki Takagi, Kaori Ota & Yoko Uchida. 2007. The intermediate degree of VOT in Japanese initial voiceless stops. *Jounal of Phonetics* 35. 439–443.

Schellinger, Sarah. 2008. *The role of intermediate productions and listener expectations on the perception of children's speech*. M.S. thesis, Department of Communicative Disorders, University of Wisconsin-Madison.

Smit, Ann Bosma, Linda Hand, J. Joseph Freilinger, John E. Bernthal & Ann Bird. 1990. The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders* 55. 779–798.

So, Lydia K. H. & Barbara Dodd. 1995. The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language* 22(3). 473–495.

Stoel-Gammon, Carol. 1996. On the Acquisition of Velars in English. In Barbara Bernhardt, John Gilbert & David Ingram (eds.), *Proceedings of the UBC International Conference on Phonological Acquisition,* 201–214. Somerville: Cascadilla Press.

Stokes, Stephanie F. & Carol Kit Sum To. 2002. Feature development in Cantonese. *Clinical Linguistics and Phonetics* 16(6). 443–459.

Trubetskoî, Nikolaî S. 1958. *Grundzüge der Phonologie*. Göttigen, Vandehoeck & Ruprecht.

Tsurutani, Chiharu. 2004. Yōon shūtoku katei ni mirareru daiiti daini gengo no on'in kōzō no eikyō [Acquisition of Yo-on (Japanese contracted sounds) in L1 and L2 phonology]. *Second Language [Journal of the Japan Second Language Association]* 3. 27–47.

Vihman, Marilyn M. 1993. Variable paths to early word production. *Journal of Phonetics* 21(1). 61–82.

Vygotsky, Lev S. 1978. *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press

Wada, Takuo, Motoyasu Yasumoto, Noriyuki Ikeoka, Yoshishige Fujiki & Ryuichi Yoshinaga. 1970. An approach for the cinefluorographic study of articulatory movements. *Cleft Palate Journal of Child Language* 7(5). 506–522.