When children first begin to produce vocalizations that listeners recognize as meaningful words of the ambient language, they show characteristic consonant errors. These misarticulations are generally perceived in terms of categorical processes that delete target consonants or that substitute one adult target consonant for another. For example, English-acquiring children are stereotypically perceived as substituting alveolar stops for velars and voiced stops for aspirated, so that a 2-year-old’s productions of *goat* and of *tub* both might be transcribed as having an initial /d/ (or [t] in a more narrow “phonetic” transcription). There are two seemingly contradictory characterizations of these misarticulations. In work such as Locke’s (1983) masterful review monograph, we see broad generalizations about typical consonant inventories and typical misarticulations, which seem to support Jakobson’s (1941) “laws of irreversible solidarity” — i.e., “implicational universals” such that mastery of more “marked” sounds implies the prior mastery of less “marked” sounds. By contrast, the detailed quantitative observations in longitudinal cross-language studies such as Vihman (1993) show an enormous variability in production patterns in young children, and at least part of this variability seems to be correlated with differences in consonant frequencies across the different lexicons.

In this paper, we examine whether these two characterizations can be reconciled by relating them to a broad-grained typological generalization about consonant inventories. Adopting Lindblom and Maddieson’s (1988) terms, the generalization is that smaller inventories tend to include only simple “basic” consonants that also occur in languages with “elaborated” or “complex” ones. For example, 413 of the 451 languages in the UPSID database (Maddieson & Precoda, 1989) are listed as having [t] whereas only 111 are listed as having the aspirated counterpart [th]. Similarly, 403 languages are listed as having [k], whereas only 13 are listed as contrasting the palatalized [kʲ] and only 3 have the “complex” stop [kʰ]. This generalization echoes the implicational universals that seem to hold in broad-grained characterizations of the relative order of consonant mastery across languages. That is, just as languages tend to incorporate consonants such as [t] and [k] before consonants such as [th], [kj], and especially [kʰ] children tend to master the basic consonants first.

Of course, children do not master consonants in isolation. Indeed, many of the most basic consonants cannot be produced audibly in isolation. The “sound” of [t] or [k] when there is no neighboring vowel or vowel-like release is a silence indistinguishable from the “sound” of [th] or [kj]. 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. A first question we ask, then, is whether the qualitative generalization about consonant inventories across languages also holds as a set of quantitative generalizations about the inventories of words in specific languages. Do basic consonants which occur in many languages also tend to occur in many words in the lexicons of these languages? Conversely, do elaborated and complex consonants that occur in few languages occur in few words in the lexicons of those languages? If the broad-grained qualitative generalization about consonant inventories also holds as a fine-grained quantitative generalization about lexicons, then this might help explain the developmental universals. That is, since mastering a consonant means learning to perceive and reproduce it in a particular set of words, we might expect a very young child to be more accurate in producing consonants that occur in many of the words of the ambient language and less accurate in producing ones that occur in few words. Insofar as these frequencies are similar across languages because of the forces of Evolutionary Phonology (Blevins, 2006), we might expect similarities to the accuracy rates and error patterns of children.

To address these issues, we compared accuracy of target consonants in a database of productions elicited in a word-repetition task from 2- and 3-year-old children acquiring Cantonese, English, Greek, or Japanese and we also compared relative frequencies of the same target consonants in the lexicons of these languages. For each language, the target consonants were all of the lingual obstruents that occur word-initially in at
least three words that a young child is likely to know. For example, in Japanese, target [t] was elicited in tamago ‘egg’, tebukuro ‘glove’, tisshu ‘tissue’, tora ‘tiger’, and 8 other words. We are interested in lingual obstruents because these are less “transparent” to the young language learner than 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 both tebukuro and tora, for example, a Japanese-acquiring child must deduce not only that the silence at the beginning of each word 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 difference in second formant transition after its release.

Accuracy rates for “shared” sounds were more or less correlated across language pairs ($r^2$ ranged from 82% for Cantonese/Greek to 2% for Japanese-English), but pair-wise correlations of frequencies missed many sounds. We therefore developed a more global measure of expected frequency by first ranking frequencies of all of the consonants that occurred in any language, counting all unattested consonants for a language as equally at the bottom of the list. We then correlated the log frequencies of the consonants within each language against the mean ranks for the three other languages. These mean ranks accounted for at least 20% of the variance in the log frequencies of each language’s inventory of lingual obstruents, although the correlation was significantly predictive only for Greek and English. We interpret this result as offering modest support to the idea that the same consonants tend to occur in many words across languages and that “marked” consonants that are attested in few languages tend to occur in fewer words in those languages.

To test the corollary idea that children are quicker to master consonants that occur in many words, we regressed the accuracy rates for the consonants in the word productions elicited from the twenty children for each language against both the log frequencies of the consonants in the target language and their mean ranks across the languages. The different components of the regression models accounted for different amounts of the variation in accuracy — e.g., in Cantonese, the global frequency ranking was the only significant component of the model, whereas in English only the language-specific consonant frequency was significant. We interpret these results in terms of historical contingencies that have conspired to make for some unexpected patterns of lexical frequencies. For example, 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 recent loanwords such as γκολ [gol] ‘goal’. Also, in both Cantonese and English, the aspirated stops are less accurate than the unaspirated stops. In Cantonese, the aspirated stops are also less frequent than the unaspirated stops, since the latter represent two Middle Chinese types (voiced and voiceless unaspirated) which merged historically. In English, by contrast, the aspirated stops are more frequent, having developed historically from the “basic” voiceless stops of common Germanic. In short, calling two sounds the “same” across different languages is an identity by analogy, not homology. Each child acquires a phonology by learning the auditory and articulatory patterns of the words of the specific ambient language. Different histories for analogous consonants can lead to different frequencies, different accuracies, and different relationships between frequency and accuracy.

Acknowledgements
Work supported by NIDCD grant 02932 to Jan Edwards.

References