Covert contrast in child phonology is not necessarily extragrammatical

Tara McAllister Byun (NYU), Peter Richtsmeier (Purdue), Edwin Maas (Univ. of Arizona)

The authors would like to recognize Sharon Inkelas (UC Berkeley) and Yvan Rose (Memorial University of Newfoundland) for their contribution to this work.

1. Background: Covert contrast in child phonology

The study of speech acquisition has a long history of debate over whether children's speech sound errors can be characterized as primarily phonetic or phonological in nature. At one extreme, it has been asserted that all child speech patterns are a reflection of young speakers' performance limitations, unrelated to their grammatical competence (e.g. Hale & Reiss, 1998, 2008). However, this "pure performance" approach has the drawback of missing the overwhelming similarities between child patterns and adult phonological processes—they are systematic, subject to language-specific variation, and sensitive to phonological structures like syllables and feet (e.g. Inkelas & Rose, 2007). At the other end of the continuum are models characterizing child and adult grammars as different permutations of a single set of formal constraints (e.g. Dinnsen et al., 2011). This position has its own drawbacks, specifically when it comes to accounting for child patterns such as positional velar fronting, which lack counterparts in adult typology (Inkelas & Rose, 2007). A common compromise between these two extremes seeks to associate different error types with different levels or domains of processing. In these models, errors of substitution and omission are thought to arise from the domain of phonology, whereas errors involving sub-phonemic distortion are considered phonetic or performance-based (see discussion in Gibbon, 1999).

This simple dichotomy faces a challenge from the phenomenon of sub-perceptible or covert contrasts in child phonology (Macken & Barton, 1980; Scobbie, 1998; Gibbon, 1999, Richtsmeier, 2010; *inter alia*). In covert contrast, the speaker produces a measurable contrast between two categories (e.g. reliably different VOT values for voiced versus voiceless stop targets), but the distinction is in some way non-canonical and therefore not perceptible under ordinary circumstances. Based on phonetic transcription, these errors would be characterized as substitutions and would thus be classified as phonological in the framework stated above. However, if it is indeed the case that these errors involve substitution of one phoneme for another, we would not expect to find a phonetic difference between supposedly neutralized categories. For this reason, covert contrast has conventionally been characterized as a phenomenon at the level of phonetic implementation: "the result of variable, gradient phonetic articulatory routines, not categorial, cognitive phonological rules" (Scobbie et al., 2000).

However, this characterization faces a challenge from recent evidence that covert contrast is more pervasive than previously realized. There is a tendency to underestimate the true prevalence of covert contrast, which can only be detected in the limited set of parameters that the experimenter undertakes to measure (Scobbie et al., 2000; Richtsmeier, 2010). Since the human ear responds to a wide array of acoustic parameters, perceptual rating tasks that require a gradient response, such as visual analog scaling (VAS), may offer a more sensitive index of covert contrast than acoustic or articulatory measurements (Munson et al., 2010). Using evidence from VAS, Munson et al. argued that covert contrast in acquisition may be "the rule rather than the exception." In light of these findings, if we continue to treat covert contrast as strictly extragrammatical, we may end up dismissing virtually all child speech patterns as performance phenomena (Hale & Reiss, 1998, 2008). This leads us to a natural question, addressed in this paper: Is covert contrast incompatible with a phonological account of child speech patterns?

We argue that a phonological analysis of covert contrast is possible in a model that posits *grammatical pressure to reuse a previous output* (Becker & Tessier, 2011; Tessier, 2012; McAllister Byun, Inkelas, & Rose, 2012). In this class of model, an error might originally arise as a consequence of performance factors, then continue to be produced under the influence of a phonological constraint such as USELISTEDERROR (Tessier, 2010) or RECYCLE (McAllister Byun et al., 2012). These models assume exemplar-based storage of input and output forms previously experienced by the speaker (e.g. Johnson, 1997; Pierrehumbert, 2001); phonemic categories are defined in terms of probability distributions over these clouds of episodic traces. To capture the systematic nature of children's error patterns, it is assumed that exemplar space interfaces with a formal grammar that can be represented with ranked or weighted constraints. McAllister Byun et al. (2012) proposed that conventional feature-based grammars are enhanced by two phonetically-sensitive constraints. The faithfulness constraint PMATCH favors a candidate that represents the closest perceptual match for the adult acoustic target, while the markedness constraint RECYCLE favors a candidate whose associated motor plan can be realized with a high degree of reliability—independent of the perceptual/acoustic accuracy of the resulting output. The balance between RECYCLE and PMATCH is negotiated within the grammar, but any previous production can be a candidate; it need not be a canonical example of some other phoneme category. This provides an explanation as to why children's errors may feature phonetically intermediate forms yet still respond primarily to phonologically defined conditioning influences. This model can also account for the well-attested phenomenon of U-shaped curves in developmental phonology, in which children initially exhibit variable output, including some correct productions, before proceeding to a stage of systematic incorrect production. While the details of the mechanism are outside of the scope of this paper, the basic intuition is that the child might initially attempt an adult-like output but make so many performance errors that the grammar enforces systematic production of the error form. By this logic, the U-shaped learning curve can be viewed as a diagnostic for phonologization of an error pattern. As a preliminary test of the hypothesis that covert contrast can arise through a process of phonologization of performance errors, we reviewed previous literature for reports of cases of U-shaped learning in which the stabilized error form features covert contrast. A relevant example was identified in a study of the acquisition of initial clusters and voicing contrasts by Catts & Kamhi (1984), also discussed in Scobbie (1998).

2. Data from Catts & Kamhi (1984)

Because this data set was collected many years ago, some information (e.g. measures of dispersion) is missing and cannot be retrieved. However, the data are adequate for the purpose of this proof-of-concept study. In their investigation of the development of voicing and clusters, Catts & Kamhi recorded six typically developing children once per month over a period of at least five months. Children, who ranged in age from 1;9 to 2;10 at the start of the study, produced single words in a naming task. The target words contained initial s-stop clusters or initial singleton stops. A child's participation was terminated when onset clusters were produced with > 90% accuracy. Catts & Kamhi measured VOT durations for all stops in singleton, reduced cluster, and unreduced cluster contexts. They generally did not report the actual VOT durations, but instead reported the percentage of stops that fell into long-lag, short-lag, and prevoiced categories. However, the case that we focus on, child AS, was reported with more detailed VOT data. AS was followed from age 1;9 to 2;5. Unlike other participants, she only reached approximately 60% correct production of clusters before she withdrew from the study due to illness. Catts & Kamhi reported that in sessions 1-3 (ages 1;9-1;11), AS produced both voiced and voiceless targets predominantly with short-lag VOT values. All s-clusters were reduced to the stop component. Of these, 80% were realized with short-lag VOT values, while 12-16% were prevoiced. In sessions 4-8 (ages 2;0-2;5), AS began to realize initial singleton stops with VOT values appropriate for the target voicing specification. In this period, /sp/ and /sk/ clusters continued to undergo reduction to stops, while /st/ clusters were reduced to /s/. AS's VOT values for cluster-reduced stops were more variable during this period. From session 4 to session 7, 47-53% of her reduced clusters featured short-lag VOT, while 32-47% showed long-lag VOT. Session 8 featured 100% long-lag VOT values, but only 2 tokens of clusters reduced to stops were observed in this session. Changes in percent occurrence of VOT ranges over time are depicted in Figure 1.

For AS's sessions 4-8 only, Catts & Kamhi reported mean VOT values for cluster-reduced stops, voiced singleton stops, and voiceless singleton stops, plotted in Figure 2. They found that the mean VOT value for voiced singleton targets (21.3 ms) was shorter than the mean for stops resulting from cluster reduction (39.7 ms), which in turn was shorter than mean for voiceless singleton targets (64.4 ms). These differences were statistically significant (voiced singleton versus reduced cluster: $t(103) = 4.0$, $p < .01$); reduced cluster versus voiceless singleton; $t(101) = 4.2$, $p < .01$). It is not reported whether the cluster-reduced stops, which featured intermediate VOT values, were transcribed as voiced, voiceless, or were assigned variably to both categories. Regardless of the direction of the perceptual neutralization, though, the fact that a measurable distinction was maintained between all three categories indicates that the data from AS can be analyzed as a case of covert contrast.

Figure 1. VOT in cluster-reduced stops produced by AS over time. From Catts & Kamhi (1984).

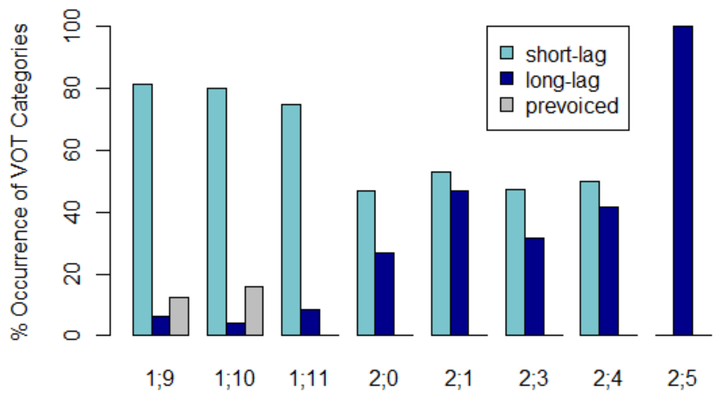
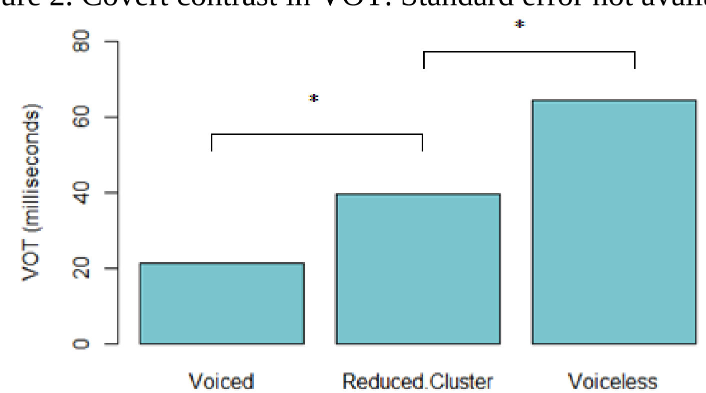


Figure 2. Covert contrast in VOT. Standard error not available from original source.



3. Analysis: Covert contrast in a U-shaped learning trajectory

The VOT data from AS are of interest because there is an apparent reversal in the overall trajectory of increasing faithfulness to the adult target. As Figure 1 shows, AS initially realized reduced clusters with short-lag VOT, then changed to long-lag VOT. Since adult clusters feature short-lag VOT, this change in AS's output represents a shift to a form that is phonetically a poorer match for the adult target. This change also does not receive a ready explanation from the standpoint of ease of articulation. Long-lag VOT requires more precise coordination of laryngeal and articulatory gestures than short-lag VOT, an explanation invoked to explain why short-lag VOT emerges earlier in typical development. So if AS were merely producing the form that was easiest from a motor control standpoint, we would expect her to continue using short-lag VOT in her reduced clusters. This example of U-shaped learning is particularly noteworthy *because it occurs in the context of covert contrast in VOT*: when AS's realization of cluster-reduced stops shifted to a less adult-like VOT range in sessions 4-8, the distribution of VOT values for cluster-reduced stops was statistically distinct from both voiced and voiceless singleton stops.

What motivation can we offer for AS's shift toward less accurate long-lag stops for reduced cluster targets in sessions 4-8? Scobbie (1998) proposed that the phenomenon may arise at the level of the phonetics-phonology interface, which "assigns aspiration directly as the interpretation of the /s/ in /st/ rather than assigning a lingual fricative gesture." That is, although long-lag VOT in a reduced cluster is gesturally less similar to the adult target than short-lag VOT, the aspiration noise may serve as a perceptual substitute for the frication noise of the /s/, which the child cannot yet realize consistently. However, Scobbie (1998) does not propose a mechanism for formal implementation of this insight. We suggest that the RECYCLE model (McAllister Byun et al., 2012) can fill this need, with the following logic. In AS's sessions 1-3, motor limitations made the long-lag stop an unstable target that incurred frequent performance errors. The constraint RECYCLE thus exerted a grammatical influence favoring production of the simpler and more stable short-lag stop. By sessions 4-8, AS had experienced some refinement in motor control, which allowed her to produce long-lag stops with acceptable (though still not perfect) consistency. AS also perceived that the aspiration noise of a long-lag stop made it a better match for an s-stop cluster than the short-lag stop on its own; this perceptual similarity is favored grammatically by PMATCH. As long as motor execution is sufficiently reliable (i.e., the RECYCLE violation is not too high), a long-lag candidate will be favored for the realization of clusters. Covert contrast can arise in this situation because the perceptual target is defined by the noisiness of the release/aspiration, not by the time lag between stop release and vowel onset. Because voiceless stops are not assumed to be the target, there is no reason to expect reduced clusters to conform to the phonetic properties of that category. In summary, the case of AS provides preliminary evidence that covert contrast can occur in the context of U-shaped learning, considered a diagnostic for phonologization. This pattern can be modeled economically in a constraint-based grammar that is supplemented by exemplar-based information about the perceptual similarity and motor stability of candidate forms.

Becker, M., & Tessier, A.-M. (2011). Trajectories of faithfulness in child-specific phonology. *Phonology*, 28, 163-196.

Catts, H.W., & Kamhi, A.G. (1984). A simplification of /s/ + stop consonant clusters: a developmental perspective. *Journal of Speech and Hearing Research*, 27, 556-561.

Dinnsen, D.A., Green, C.R., Morrisette, M.L., & Gierut, J.A. (2011). On the tnteraction of velar fronting and labial harmony. *Clinical Linguistics & Phonetics*, 25, 231-251.

Gibbon, F. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research* ***42***, 382–97.

Hale, M. & Reiss, C. (1998). Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry*, 29, 656–83.

Hale, M. & Reiss, C. (2008). *The phonological enterprise*. Oxford: Oxford University Press.

Inkelas, S. & Rose, Y. (2007). Positional neutralization: a case study from child language. *Language* ***83***, 707-36.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (eds), Talker variability in speech processing, 145-165. San Diego: Academic Press.

Macken, M. A., & Barton, D. (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, 41–74

McAllister Byun, T., Inkelas, S., & Rose, Y. (2012). Transient Phonology, CON and Child Phonological Processes. Paper presented at the 20th Manchester Phonology Meeting, Manchester, UK.

Munson, B., Edwards, J., Schellinger, S., Beckman, M. E., & Meyer, M. (2010). Deconstructing phonetic transcription. *Clinical Linguistics and Phonetics*, 24, 245-260.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. L. Bybee & P. Hopper (eds), Frequency and the emergence of linguistic structure, 137-157. Amsterdam: John Benjamins.

Richtsmeier, P. T. (2010). Child phoneme errors are not substitutions. Toronto Working Papers in Linguistics, 33. Available at http://twpl.library.utoronto.ca/index.php/twpl/article/view/6889

Scobbie, J. M. (1998) Interactions between the acquisition of phonetics and phonology. In Gruber et al. (eds), *Papers from the 34th Annual Regional Meeting of the Chicago Linguistic Society*, Volume II, 343-358. Chicago: Chicago Linguistics Society.

Scobbie, J., Gibbon, F., Hardcastle, W., & Fletcher, P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In M. Broe & J. Pierrehumbert (eds), *Papers in Laboratory Phonology V*, 194-207. Cambridge, UK: Cambridge University Press.

Tessier, A.-M. (2012). USELISTEDERROR: a grammatical account of lexical exceptions in phonological acquisition. In S. Lima, K. Mullins, & B. Smith (eds), Proceedings of NELS39, Volume 2, 813-827. Amherst, MA: GLSA.