Published in final edited form as:

Lingua. 2015 June 1; 160: 38–53. doi:10.1016/j.lingua.2015.02.007.

Perceptual discrimination across contexts and contrasts in preschool-aged children

Tara McALLISTER BYUN [Assistant Professor]

Department of Communicative Sciences and Disorders New York University Steinhardt School of Culture, Education, and Human Development

Abstract

This paper investigates a proposed phonetically-based account of developmental phonological patterns that lack counterparts in adult typology. Adult listeners perceive some phonemic contrasts more accurately than others, and these differences in perceptual recoverability are posited to represent one influence on phonological typology. One hypothesis suggests that children and adults could differ in their patterns of relative perceptual sensitivity, and these differences could form the basis for some child-specific phonological patterns in production. However, there has been a lack of empirical evidence to support this claim. This study used a nonword discrimination task to investigate differences in perceptual recoverability across contrasts and contexts in typically-developing preschool children. Participants heard nonwords that were identical or differed by a single segment in initial or final position. Results revealed general agreement between child and adult listeners in the relative discriminability of different featural contrasts. For certain contrasts, discrimination accuracy was significantly greater in initial than final position, mirroring an asymmetry seen in adults. Overall, these results suggest that perceptual discrimination in preschool-aged children is broadly congruent with patterns of relative sensitivity observed in adult listeners. These findings suggest that factors other than perceptual recoverability should be explored to account for child-specific phonological patterns.

Keywords

speech perception; perceptual development; perceptual recoverability; speech acquisition

1. Introduction

1.1 Phonetic pressures as an influence on phonological typology

A common assumption in the study of language acquisition is that of *continuity* between child and adult grammars: that is, the rules or principles governing a child's linguistic behavior should not pose an exception to generalizations that are universal across adult

Corresponding author: Tara McAllister Byun, New York University Department of Communicative Sciences and Disorders, 665 Broadway, Suite 900, New York, New York USA 10012 tara.byun@nyu.edu Tel: 1 212 992 9445 Fax: 1 212 995 4356.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

languages (e.g., Macnamara, 1982; Pinker, 1984). In the context of developmental phonology, the continuity hypothesis is challenged by the existence of *child-specific phonological patterns*, which are attested or even common in child speech but find no counterpart in adult typology. For example, many children acquiring English exhibit a pattern of consonant harmony in which a coronal consonant assimilates to a labial or velar consonant across an intervening vowel (e.g. [gɔg] for “dog,” [kʌk] for “cup”), whereas no adult grammar permits long-distance assimilation for major place of articulation (e.g., Pater, 1997; Becker & Tessier, 2011). These areas of divergence between child and adult phonology have drawn interest as a potential source of insight into broader questions about the nature of phonological knowledge. Various factors have been posited to play a role in determining which sound patterns recur in cross-linguistic typology: universal markedness (e.g., Hale & Reiss, 2008), phonetic pressures (e.g. Hayes, Kirchner, & Steriade, 2004), or biases in the diachronic transmission of language (e.g. Blevins, 2004). It is probable that all of these factors interact to shape phonological typology. However, following previous work on child-specific speech patterns (e.g. Inkelas & Rose, 2007; Dinnsen & Farris-Trimble, 2008; McAllister Byun, 2011), the present work will focus on the hypothesis that phonetic factors play an important role in shaping child phonology. It is well-established that the physical experience of producing and perceiving speech differs across child and adult speakers (e.g., Kent, 1992; Nittrover, 2002). If phonological patterns can arise from phonetic pressures (e.g., Hayes et al., 2004), these phonetic differences are naturally predicted to give rise to some degree of divergence in the systematic phonological patterns exhibited by child and adult speakers. Despite the present emphasis on phonetic (especially perceptual) influences on phonology, it should be kept in mind that these factors exist in a potentially broader context of universal markedness and diachronic pressures.

1.2 Perceptual recoverability in adult speech patterns

Previous research has clearly established that phoneme contrasts within a language can differ in acoustic-perceptual salience. That is, certain sound pairs are more readily discriminated than others across a variety of listening conditions, as in the contrast between the highly confusable fricatives /f/-/θ/ and the highly distinctive fricatives /s/-/ʃ/ (e.g., Cutler, Weber, Smits, & Cooper, 2004). These differences in discriminability are largely rooted in physical properties of the acoustic signal, such as loudness and duration. For example, the frication noise generated in the production of the sibilants /s/ and /ʃ/ is considerably louder than the frication noise for nonsibilants such as /f/ and /θ/, offering a straightforward explanation for the lower discriminability of the latter contrast (e.g., Wright, 2004). Similarly, Wright (2004) found that perception of stop place contrasts was more negatively affected by increasing levels of background noise than perception of intensity-matched fricative place contrasts. The greater perceptual recoverability of fricative contrasts was attributed to their longer duration, which provides the listener with an extended window in which to detect the target contrast. This paper will use the term *cue* to refer to any acoustic information that contributes to the perceptual recoverability of a phoneme contrast, and *cue robustness* to refer to the acoustic salience of those elements¹. Some phonemes carry strong *internal cues* (i.e., they can be perceived while standing alone), while others are perceived primarily through *transitional cues* created by coarticulation with an adjacent vowel or sonorant consonant.

Many of the results of studies of perceptual cue robustness are highly intuitive, such as the finding that phonemes associated with a loud or long acoustic signal are perceived more reliably than phonemes with short or low-intensity acoustic cues. The most interesting contribution of this body of experimentation has emerged from a line of research proposing that differences in perceptual salience can contribute to a deeper understanding of the cross-linguistic typology of phonological systems. Differences in perceptual recoverability have been argued to account for implicational generalizations regarding the environments in which a given sound or contrast can appear in a language: that is, the occurrence of a given contrast in a context of low perceptual salience implies that it can also occur in a context of greater perceptual salience (e.g., Steriade, 2001; Henke, Kaisse, & Wright, 2012). For example, many languages allow sibilant fricatives to occur in a wider range of consonant cluster contexts than other obstruents. English permits a variety of onset clusters with rising sonority (e.g., /bl-, fr-, tw-, sn-/), but only sibilants can appear in onset clusters of flat or falling sonority (e.g., /st-, sp-, sk-/; cf. */pt-, fp-, rt-/). The broader distribution of sibilants is attributed to their robust internal perceptual cues, which allow the listener to recover the identity of a sibilant even when it does not occur adjacent to a high-sonority segment that can act as a carrier of transitional cues (Henke et al., 2012).

Another factor frequently studied in connection with the influence of perceptual recoverability on phonological typology, which will play a key role in the study reported below, relates to asymmetries in the distribution of consonants in syllable-initial versus syllable-final position. Cross-linguistically, CV syllables are more common than VC syllables, and within a language, phoneme contrasts that occur in final position tend to represent a subset of the contrasts that can occur in initial position (e.g., Bell & Hooper, 1978; Redford & Diehl, 1999; Beckman, 1997). For example, European Portuguese contrasts labial, alveolar, and postalveolar fricatives in onset position, but only postalveolar fricatives can appear in coda position (Mateus & d'Andrade, 2000). These phenomena have been attributed to the fact that perceptual cues to contrasts among non-sonorant consonants are carried primarily in the transition to or from an adjacent sonorant (e.g., Henke et al., 2012). Crucially, a number of experimental studies have found that transitional cues to consonant place are more perceptible in CV than VC contexts (e.g., Benkí, 2003; Dorman, Raphael, & Repp, 1975; Fujimura, Macchi, & Streeter, 1978; Miller & Nicely, 1955; Redford & Diehl, 1999; Wright, 2004). For example, Fujimura, Macchi & Streeter (1978) created VCCV disyllables with conflicting consonant place cues by splicing together VC and CV transition portions (bursts excluded) from naturally produced Japanese pseudowords. In a forced-choice task, the place of articulation consistent with cues from the CV transition was selected by Japanese-speaking listeners in 85.1% of trials and by English-speaking listeners in 84.5% of trials. The same bias was observed when stimuli were played in the reverse direction, indicating that the results did not depend on some articulatory asymmetry inherent in stimulus production. In summary, the literature offers a multiply-replicated finding that cues in syllable-initial position have intrinsically greater salience than cues in syllable-final position, and this asymmetry has been argued to play a causal role in

¹Discussion in this paper will focus on acoustic intensity and duration as the major predictors of cue robustness. However, other influences on robustness have been proposed, such as the degree of modulation or acoustic change between sounds, and relative sequencing of sounds with differing acoustic properties (see discussion in Henke, Kaisse, & Wright, 2012).

shaping the cross-linguistic tendency for languages to allow a broader range of phonemic contrasts in prevocalic than postvocalic contexts.

Although a number of studies have replicated the finding that the perceptual recoverability of consonant contrasts is enhanced in presonorant contexts, exceptions to this generalization have also been identified. The perceptual advantage for initial over final contrasts may not apply equally across all target consonants (Redford & Diehl, 1999; Wright, 2004). Other research has reported that positional asymmetries in perceptual recoverability may be influenced by listening conditions, such as the level of background noise. Cutler et al. (2004) found that at moderate to high levels of background noise, identification was more accurate for consonants in syllable-final than syllable-initial position, contrasting with the asymmetry reported in most previous literature. At the lowest noise level, which was comparable to the listening conditions used by Redford & Diehl (1999), there was no difference in listeners' accuracy for initial and final consonants. Discussion in the remainder of this paper will endorse the majority finding that initial/prevocalic position has a general tendency to enhance the perceptual recoverability of consonant place contrasts. However, it is acknowledged that this general advantage may not hold equally across all contrasts and contexts.

1.3 Perceptual recoverability in child speech patterns

A question that has received surprisingly limited attention in the literature is whether speech patterns in developmental phonology can be understood to be shaped by perceptual recoverability in the same way that has been posited for adult typology. Previous work has shown that child speakers show less acute perceptual discrimination of phonemic contrasts than adults, with ongoing refinement until 12 years of age or beyond (Hazan & Barrett, 2000). Other research has documented qualitative changes in patterns of relative response to different perceptual cues over the course of development. Most notably, Nittrouer & Studdert-Kennedy (1987) found that young children give more perceptual weight to transitional cues than older children and adults. This result has been replicated several times (e.g., Nittrouer, 2002; Mayo, Scobbie, Hewlett, & Waters, 2003), although it may hold only for certain contrasts (e.g., Mayo & Turk, 2004). Thus, it is entirely possible that there are differences between child and adult listeners in the relative perceptual recoverability of different contexts or featural contrasts.

The previous section laid out a proposed link between perception and production in adult phonological typology: languages tend to allow the maximum range of contrasts to be realized in contexts that maximize the robustness of perceptual cues, while limiting contrasts in contexts of low perceptual recoverability. There is also ample evidence that perception and production are linked in children's phonological acquisition, with a large number of studies reporting that children show specific difficulty discriminating the same contrasts that they neutralize in production (e.g., Hoffman, Daniloff, Bengoa, & Schuckers, 1985; Locke, 1980; McAllister Byun, 2012a; Rvachew & Jamieson, 1989; Velleman, 1988; Whitehill, Francis, & Ching, 2003; Ygual-Fernández & Cervera-Mérida, 2013; but cf. Gierut, 2006). By analogy to the adult case, we can predict that error patterns in child production should be more likely to affect contrasts of low perceptual salience than contrasts with robust cues.

That is, just as patterns that neutralize contrast in adult phonology are more likely to apply in contexts of low perceptual salience (e.g., neutralization across places of articulation in coda fricatives in European Portuguese), we might expect child patterns that systematically neutralize adult phonological contrasts to apply more strongly in a context of low perceptual salience than in a highly salient context.

However, some child production patterns are unexpected from the standpoint of perceptual recoverability in adults. A noteworthy example is the literature on “neutralization in strong position” in developmental phonology, i.e. child-specific patterns that neutralize phonemic contrast selectively in word- or foot-initial contexts. The best-known example is positional velar fronting, in which alveolar and velar place are neutralized in initial but not final contexts (Bills & Golston, 2002; Chiat, 1983; Dinnsen, 2008; Inkelas & Rose, 2003, 2007; McAllister Byun, 2012b; Morrisette, Dinnsen, & Gierut, 2003; Stoel-Gammon, 1996). Patterns of initial gliding, stopping, and deaffrication have also been reported (Dinnsen & Farris-Trimble, 2008; Inkelas & Rose, 2007; McAllister Byun, 2011). Examples of initial velar fronting and initial fricative gliding are provided in (1). By limiting the range of contrasts that occur in initial position while preserving a broader range of contrasts word-finally, these patterns violate a strong cross-linguistic generalization thought to be grounded in perceptual recoverability.

(1) Child phonological patterns affecting word-initial but not word-final position (data from “Ben,” aged 3;10 - 4;3; McAllister Byun, 2012a).

a. Fronting of velar stops in initial but not final position	
[daut]	“card”
[dos]	“ghost”
[wʊk]	“look”
[dak]	“cake”
b. Gliding of fricatives in initial but not final position	
[joɪ]	“sewing”
[jiba]	“zebra”
[mas]	“mouse”
[babajis]	“strawberries”

How can we account for this discrepancy between child and adult patterns of positional neutralization? One possibility is that children and adults experience the same asymmetries in perceptual recoverability, and child-specific patterns of positional neutralization are the product of factors external to perception. It has been proposed that child-specific patterns of positional neutralization are driven by licensing requirements or representational constraints (see discussion in Goad, 2004). An alternative account holds that these patterns reflect constraints on articulation: Inkelas & Rose (2003, 2007) and McAllister Byun (2012b) have argued that the larger magnitude of articulatory gestures in prosodically strong positions presents a more challenging motor control task, leading to systematic error patterns that preferentially affect initial position. The details of these proposals fall outside of the scope

of the present paper. Instead, this study was devised to test a third hypothesis that has been advanced in previous literature on child-specific phonology, as discussed below.

The third alternative is that children's production typology is indeed shaped by differences in perceptual recoverability, but adults differ with respect to what can be considered a perceptually prominent environment. An account along these lines has been proposed by Dinnsen & Farris-Trimble (2008, 2009). Framing their account in the formalism of Optimality Theory (Prince & Smolensky, 2004), Dinnsen & Farris-Trimble proposed that prominence is determined by the ranking of two violable constraints termed *INITIALPROM* ('The initial constituent of a syllable/foot/prosodic word must be prominent') and *FINALPROM* ('The final constituent of a syllable/foot/prosodic word must be prominent'). When *INITIALPROM* dominates *FINALPROM*, speakers treat syllable-initial position as more prominent than final; with the reverse ranking, final positions are prominent.² A faithfulness constraint ID-Prom is then proposed to capture speakers' tendency to selectively preserve contrast in whichever context is currently classified as prominent. Dinnsen and Farris-Trimble posited that *FINALPROM* dominates *INITIALPROM* in the initial, default constraint ranking of a child's phonology, but *FINALPROM* is necessarily demoted as children undergo a shift in perceptual sensitivity over the course of development. In this account, child patterns of neutralization in initial position and adult patterns of neutralization in final position are comparable in the sense that both preferentially affect contrasts in the context that is least salient to the speaker at the time.

However, literature to date has offered little indication that child speakers demonstrate enhanced perception of final relative to initial contrasts. Dinnsen & Farris-Trimble cite examples (Brooks & MacWhinney, 2000; Coady & Aslin, 2004) providing evidence that the entirety of the rhyme (nucleus and coda) has greater salience to young child listeners than individual segments in onset position. However, child patterns of neutralization in strong position in production involve selective preservation of contrast in individual segments in coda relative to onset position. The studies cited by Dinnsen & Farris-Trimble (2008) do not offer evidence that children in the same age range would discriminate segmental contrasts more acutely in coda than onset position. In fact, several studies of infant perception have reported greater sensitivity to contrasts in onset relative to coda position (Juszyk, Goodman & Bauman, 1999; Zamuner, 2006; Altvater-Mackensen & Fikkert, 2010). For children in the preschool age range, there is very little systematic evidence about differences in recoverability across contexts or contrasts.

To address this question, the present study collected nonword discrimination data from typically developing preschool-aged children. By comparing detection of consonant contrasts in word-initial versus word-final position, the experiment evaluated whether child patterns of perceptual sensitivity diverge from the frequently reported adult pattern in which contrasts are recovered more accurately in prevocalic than postvocalic contexts. The study

²Dinnsen & Farris-Trimble do not explicitly state that there is a developmental shift in *perceptual* prominence; it is possible that they are referring to a shift in prominence that does not encompass perception. However, they do include perceptual salience in their definition of the abstract notion of prominence: "contexts...judged to be strong, perceptually salient, or prominent" (2009, p. 113). Further, if one were to posit that positional prominence in children is unrelated to perception, the result is an unexplained discontinuity between the child and adult systems. That is, a child with *FINALPROM* >> *INITIALPROM* would show enhanced production, but not enhanced perception, of final contrasts; this same child would mature into an adult with *INITIALPROM* >> *FINALPROM* who shows enhancement with respect to both perception and production of contrasts in prominent contexts.

also compared featural contrasts such as stop versus fricative manner and velar versus alveolar place, to test whether the relative discriminability of these contrasts differs across child and adult listeners. If areas of systematic divergence between child and adult perception emerge, these differences might be invoked to account for error patterns in developmental phonology that lack counterparts in adult typology, such as preferential neutralization in initial position. By contrast, if there is general convergence between child and adult patterns of perceptual sensitivity, we should look to other domains for an explanation of child-specific speech patterns.

2. Method

2.1 Participants

Participants were monolingual English speakers recruited from a preschool in Montclair, NJ. All participants were required to pass a hearing screening at 1000, 2000, and 4000 Hz at 20 dB HL. Participants were excluded from the present study if parent report indicated a history of significant speech-language delay or disorder. A total of 18 children aged 3;6 to 5;7 (mean age 4;9) were initially enrolled in the study. Three children were excluded because they did not pass an initial screening test for comprehension of the experimental task. Analyses in this paper are based on the remaining 15 participants. The final group was made up of 9 females and 6 males. The age range remained the same, but the average age increased slightly to 4;10.

2.2 Procedure

A nonword discrimination task was created to probe perception of selected phonemic contrasts in both initial and final position. Participating children heard digitally recorded naturally produced pairs of monosyllabic CVC nonwords embedded in a carrier-phrase context ("I can say X. I can say Y.") These pairs were either identical or differed by a single segment in initial or final position. Participants used a button-press response to indicate whether the nonwords they heard were the same or different.

Studies of perceptual recoverability in adults typically involve presenting spoken syllables in differing levels of background noise and asking participants to generate or select a written response representing the syllable they heard. Because this design cannot be implemented with preliterate participants, the same-different task described here was adopted instead. However, the low task difficulty of minimal pair discrimination introduces a risk of ceiling effects. The inclusion of the carrier phrase, which obligates the listener to hold the target nonwords in working memory over intervening phonological content, was a manipulation of task difficulty intended to increase the likelihood of obtaining non-ceiling-level scores that would make it possible to observe differences across categories.

2.3 Stimuli

Experimental stimuli consisted of 48 pairs of monosyllabic nonwords³ that differed by a single phoneme in initial or final position. In addition, 46 pairs of identical nonwords were created to serve as catch trials, which prevent the participant from developing a strategy of identifying all stimulus pairs as different (Elliott & Hammer, 1988). The use of nonword rather than real word stimuli allowed for free manipulation of the positional and featural properties of target contrasts while minimizing the confounding influence of lexical frequency and familiarity. The minimal pair stimuli were designed to assess the relative discriminability of four featural contrasts that are frequently neutralized in child production patterns and also differ in cue robustness. The contrasts tested were alveolar versus velar place for stops (/t-k/ or /d-g/, henceforth T-K), alveolar versus postalveolar place for sibilants (/s-ʃ/, henceforth S-SH), fricative versus stop manner for labial place of articulation (/f-p/ or /b-v/, henceforth F-P), and fricative versus stop manner for alveolar place of articulation (/s-t/ or /z-d/, henceforth S-T).⁴ Table 1 provides sample stimulus pairs for each of the three experimental categories and the catch trial (no contrast) category. Complete stimulus materials are supplied in Appendix A.

Pilot testing, described in detail below, revealed ceiling-level accuracy among adult listeners performing the experimental task used in this study. Therefore, predictions for preschool children's patterns of relative accuracy were generated not from adult listeners' performance on these specific stimuli, but from converging evidence across previous perceptibility studies whose design and stimuli differed in various ways. The relative perceptual salience of contrasts in initial versus final position was evaluated with reference to previous findings that adult listeners recover perceptual contrasts more accurately in prevocalic than postvocalic contexts (Dorman et al., 1975; Fujimura et al., 1978; Redford & Diehl, 1999). The relative perceptibility of the selected featural contrasts was estimated from results reported in Cutler et al. (2004). Cutler et al. recorded CV and VC strings representing nearly all possible combinations in American English and presented them in three levels of background noise in a self-paced forced-choice listening task. Results were reported in the form of confusion matrices. For the analysis of featural contrasts, data were drawn from the confusion matrix in which responses from 16 monolingual English listeners were collapsed across CV and VC contexts and three signal-to-noise ratios (Cutler et al., 2004, online supplement).⁵ Responses representing confusion of a given contrast were tallied; for example, confusion of the T-K contrast was indicated by responses of /k/ for target /t/, /t/ for /k/, /d/ for /g/, and /g/ for /d/. (For consistency with the stimulus pairs used, only errors

³A small subset of items were real words that were attested 1 or 0 times in a corpus of child-directed speech (Demuth, Culbertson, & Alter, 2006).

⁴The F-P and S-T categories were originally intended to serve as a single category representing the fricative-stop manner contrast; therefore, these categories are each represented by 8 tokens, versus 16 each for the categories T-K and S-SH. However, studies of perceptual confusion in adults (e.g., Cutler et al., 2004) make it plain that the fricative-stop manner contrast does not behave equivalently across sibilant and non-sibilant fricatives. For this reason, F-P and S-T categories were analyzed separately throughout the present study.

⁵It is a limitation that the background noise level in the present study was not matched to a specific background noise condition in Cutler et al. (2004), since the features that signal consonant contrasts can be affected in different ways by different levels and types of background noise (e.g. Cooke, García Lecumberri, Scharenborg, & Van Dommelen, 2010). For the subset of contrasts tested here, though, there is a major division between more perceptible contrasts involving sibilant fricatives, and less perceptible contrasts with no sibilants. Because Cooke et al. (2010) found that transmission of the sibilant feature was similarly robust across all noise backgrounds, this major contrast can legitimately be examined in the absence of systematic manipulation of background noise.

that matched for voicing were included in the tally.) To correct for any influence of listener bias, such as a general tendency to report hearing high-frequency phonemes, the raw numbers from the Cutler et al. matrices were converted to perceptual distances using Luce's (1963) biased choice model (see discussion in McAllister Byun, 2012a).⁶ The greatest perceptual distance was calculated for the S-T contrast ($d = 4.72$), followed by the S-SH contrast ($d = 3.91$). Perceptual distance scores were lower for the T-K contrast ($d = 2.95$) and lowest of all for the FP contrast ($d = 2.09$). If perceptibility is instead calculated by raw frequency of confusions, the same rank order of contrasts is obtained, with the exception that the S-T and S-SH contrasts do not differ in confusability (F-P: 393 confusions; T-K: 209 confusions; S-SH 49 confusions; S-T: 49 confusions). These findings are consistent with expectations based on the robustness of the underlying acoustic cues: contrasts involving /s/ and/or /ʃ/ are of high salience due to the high intensity and long duration of the sibilant fricatives, while the labial stop versus fricative contrast is less perceptible due to the lower intensity of the non-sibilant fricative. The contrast between velar and alveolar stops is also of relatively low salience due to the weak nature of internal cues to stop place.

Stimulus categories were balanced for neighborhood density using Storkel & Hoover's (2010) online calculator drawing on child speech corpora. In the final set of stimuli, there were no significant differences in number of phonological neighbors between experimental and catch trial stimuli, between stimuli testing initial versus final contrasts, or between any of the four featural contrasts tested (see complete results in Appendix B). Stimulus categories were also balanced with respect to vowel context⁷ and place and manner of the non-target consonant. Finally, categories were balanced with respect to voicing of the target consonants.⁸ The single exception was the S-SH contrast, where only voiceless targets were used due to the low frequency of occurrence of the voiced postalveolar fricative /ʒ/ in English.

One parameter not taken into consideration during stimulus creation was phonotactic probability, as measured by positional segment frequency or positional biphone frequency. These parameters were analyzed post hoc using Storkel & Hoover's (2010) calculator. Comparisons revealed no significant differences between experimental and catch trial stimuli or among the four featural contrast categories, but the sets of stimuli testing contrasts

⁶For stimulus categories s_i and s_j and response categories r_i and r_j , the perceptual distance d_{ij} is calculated according to the following formula (Luce, 1963):

$$d_{ij} = -\ln \sqrt{\frac{p(r_j|s_i)p(r_i|s_j)}{p(r_i|s_i)p(r_j|s_j)}}$$

⁷The decision to use a balanced set of vowels across stimulus sets, rather than systematically manipulating vowel context for inclusion in the statistical model, was based on the following considerations. (1) Because the effect of vowel context on the perceptibility of consonant contrasts is likely to be mediated by the shape and magnitude of formant transitions, an interaction of contrast and vowel context is more likely than a main effect of vowel context. Given the small number of stimuli per contrast, it is unlikely that an interpretable interaction would be detected. (2) The strategy of repeatedly using a single vowel context or small number of vowel contexts was avoided because it would make it difficult to control for positional segment probability and biphone frequency across stimulus categories.

⁸Voicing was not included as a predictor in the statistical model because it was not a primary factor of interest in the present comparison between child and adult perception. Post hoc analysis of the data collected indicated that voiced and voiceless pairs were discriminated with very similar overall accuracy levels (voiced mean = 61%, SE = 3.1; voiceless mean = 65%, SE = 2.2).

in initial versus final position differed significantly with respect to mean positional segment frequency ($t = 3.8$, $df = 80$, $p < .01$) and mean biphone frequency ($t = 2.2$, $df = 94$, $p = .03$). Therefore, these properties were included as covariates in the analyses reported below.

Stimuli were recorded by a female adult native speaker of English (mid-Atlantic dialect) who was coached to use a slow rate of speech, a child-directed intonation contour, and to release final consonants. Target nonwords were produced in the carrier phrase “I can say X.” During stimulus recording, items were elicited via randomly ordered lists in English orthography. Catch trial items were included twice, so that the two iterations of the target nonword were distinct utterances rather than two copies of the same recording. The complete list was recorded twice in different orders with a 44.1 kHz sampling rate and 16 bit audio sample size. Using Praat (Boersma & Weenink, 2010), stimuli were rms-equalized and standardized to contain 50-ms silent intervals before and after the target utterance. Total duration and mean intensity were measured for each file, and any outliers greater than two standard deviations from the group mean were discarded. Three listeners independently rated each pair of files to identify any items that contained artifacts or were poorly matched with respect to suprasegmental properties. Problematic items were re-recorded until the listeners reached consensus that all pairs were well-matched. Stimuli were divided into two lists, each containing 47 unique pairs, to be presented in separate testing sessions. Items were pseudorandomly assigned such that each list contained an equal number of files representing each stimulus category.

The adequacy of the experimental stimuli was tested with a pilot group of 10 young adults with typical speech, language, and hearing. The adult listeners followed the same protocol described below, with the exception that they had the option to complete both stimulus lists within a single session rather than across multiple days. Unsurprisingly, adults demonstrated virtually ceiling-level performance on this task. A logistic mixed model structured as described in Section 2.5 found no significant effect of position, stimulus category, mean positional segment frequency, or mean biphone frequency, with no significant interactions. However, nonsignificant numerical differences were consistent with the direction predicted on the basis of previous adult studies. Specifically, discrimination of initial contrasts (99.6%, standard error of the proportion (SEp) = 0.4) was numerically more accurate than discrimination of final contrasts (95.4%, SEp = 1.3). In addition, there was a numerical difference in discrimination accuracy between the contrasts predicted to have lower perceptual salience (T-K and F-P: 95%, SEp = 1.4) and the more salient contrasts (S-T and S-SH: 100%, SEp = 0).

2.4 Stimulus presentation

Testing took place in a quiet room in the participants’ preschool. The experiment was presented using E-prime (Empirisoft) on a Dell laptop operating Windows XP. Audio files were presented in sound field at a loud but comfortable volume. In each trial, a fixation screen was presented for 500 ms, followed by presentation of a randomly selected pair of nonwords separated by a 1 second inter-stimulus interval; the participant’s response was then recorded. Discrimination trials were presented in blocks of five, with brief videos between blocks to serve as a break and reward.

Participants gave their responses using a customized button box (Empirisoft) with one red and one blue button. The red button, used to indicate “same,” was accompanied by a picture of two identical faces, while the blue button was accompanied by a picture of two non-identical faces. Before experimental stimuli were presented, several measures were taken to ensure that children were able to comprehend the task. Each session began with a series of images accompanied by a spoken script in which the experimenter discussed the concept of *same* versus *different*. Participants then saw a series of images of pairs of identical or different objects, and they were cued to press the red or blue button to classify each picture pair. Participants subsequently heard examples of pairs of real words that were identical or different, followed by practice trials featuring pairs of nonwords. Unlike the minimal pairs of nonwords used in the test phase, these items differed by multiple segments. Feedback was provided during all practice trials. To be included in the study, participants were required to give at least 4 correct responses in a block of 5 practice trials. If a participant failed to meet the screening criterion in 3 blocks of 5 trials, perceptual testing was discontinued. Three participants were ruled out in this fashion.

2.5 Analyses

For initial processing, participants were ranked by overall percentage of correct responses, including both correct detection of contrast and correct “same” responses to catch trial pairs. Because experiment-wide accuracy varied widely across participants (see *Results*), individual participant data were further analyzed using the exact binomial test for count data. Both catch trial pairs and experimental pairs were included in this analysis. Participants were classified in a binary fashion reflecting whether the two-tailed binomial test returned a probability significantly different from chance (i.e., 0.5) at an α level of .05. The rationale for classifying participants in this fashion, and for retaining data from participants whose performance did not differ significantly from chance, is discussed at the end of this section.

A logistic mixed model was used to analyze the influence of stimulus properties, such as the identity and position of the featural contrast, on perceptual discrimination performance. The dependent variable was accuracy in detecting contrast when a contrast was present, i.e. in experimental trials. A random effect of subject was included to adjust for any individual-level influences on performance not captured by the independent variables included in the model. The fixed effects of interest were the syllable position of the phoneme contrast (initial or final) and the identity of the contrast (T-K, S-SH, S-T, or F-P). A first-order interaction of position and contrast was included based on previous findings that the effect of position on perceptual recoverability may vary across phonemes (e.g., Redford & Diehl, 1999; Wright, 2004). A fixed effect of session was included to control for any changes as participants became more familiar with the task over the course of the experiment. Fixed effects of mean positional segment frequency and mean biphone frequency were also included. As noted above, it was necessary to control for these factors in order to interpret the effect of syllable position on perceptual discrimination, since the sets of stimuli testing contrasts in initial versus final position were found to differ with respect to these properties.

Finally, to control for the influence of overall performance on individual patterns of relative response to different types of stimuli, the model included a binary variable reflecting

whether or not the exact binomial test indicated above-chance performance at an experiment-wide level for each participant. First-order interactions of this factor with position and featural contrast were also included. This variable was necessary because the poor overall performance exhibited by children whose accuracy did not exceed chance level is compatible with several different interpretations. All of these participants demonstrated comprehension of the task by scoring at least 4/5 correct on the initial screening measure that evaluated perception of non-minimal contrasts. Thus, it is possible that their low accuracy was a genuine reflection of a lack of perceptual discrimination of minimal pair stimuli. However, participants may have achieved their passing score on the screening measure by chance, or they may have understood the task but failed to maintain adequate attention during test trials. In light of this uncertainty, it is not appropriate to exclude data from children with chance-level performance, but it also is potentially misleading to treat these results as equivalent to the data from children who performed significantly above chance. Therefore, results from all participants were included in the logistic mixed model, and the binary factor reflecting the outcome of the exact binomial test (chance- versus above-chance-level performance) was included as a fixed effect. This variable is of interest specifically in connection with its potential for interaction with the fixed effects of position and/or featural contrast. Such an interaction would indicate that the pattern of relative discrimination accuracy across different contrasts or contexts differed between children with overall above-chance versus chance-level performance.

3. Results

3.1 Overall performance

For the 15 participants who met all inclusionary criteria, Table 2 reports the hit rate (percentage of correct responses for experimental trials), false alarm rate (percentage of “different” responses to catch trials), and percent correct for all stimuli combined, including both catch trials and experimental items.⁹ The *p*-value returned by the exact binomial test is also reported. Table 2 shows that for five participants, experiment-wide accuracy did not differ significantly from chance, whereas the remaining ten participants demonstrated above-chance performance. All of the participants who performed above chance were aged 4;6 or older. However, the group with chance-level accuracy was not specific to the youngest children, including individuals as old as 5;5. In this sample, there was no significant correlation between age in months and overall percent correct (Spearman's $\rho = .17$, $p = .54$).

3.2 Regression results

In the logistic mixed model, featural contrast was a significant predictor of discrimination accuracy (T-K versus S-T: $\beta = -1.5$, $z = -2.9$, $p < 0.01$; F-P versus S-T: $\beta = -1.3$, $z = -2.2$, $p = 0.03$). The negative coefficients indicate that correct discrimination was less likely for the T-K and F-P contrasts than the S-T contrast, which served as the reference level. The other level of the featural contrast factor, S-SH, did not differ significantly from the reference

⁹Note that no participant recorded a false alarm rate substantially greater than 50%, which would be suggestive of a strategy of responding “different” to most or all items.

level ($\beta = -.71$, $z = -1.4$, $p = 0.16$). Unsurprisingly, the binary variable reflecting the outcome of the exact binomial test was also a significant predictor of discrimination accuracy ($\beta = 2.3$; $z = 3.4$; $p < 0.001$); it additionally participated in a marginally significant interaction with the factor of position ($\beta = .74$; $z = 1.9$; $p = 0.05$). There was no significant main effect of position, positional segment frequency, or positional biphone frequency. However, there was a significant interaction between position and featural contrast (for the T-K contrast, $\beta = 1.3$, $z = 2.2$, $p = 0.03$; for the F-P contrast, $\beta = 1.5$, $z = 2.3$, $p = 0.02$; the difference from the reference level was not significant for the S-SH contrast, $\beta = 1.0$, $z = 1.7$, $p = 0.09$). Finally, there was a marginally significant effect of session, reflecting a slight tendency toward improved discrimination after the first session ($\beta = .37$; $z = 1.9$; $p = 0.05$). There were no other significant effects or interactions in the model. Complete results of the regression are reported in Appendix C.

3.3 Detailed examination of significant effects and interactions

With data pooled across participants and positions, the percentage of correct discrimination was highest for the S-T contrast (73.8%, standard error of the proportion (SEp) = 4.0), which served as the reference level in the regression reported above. The S-SH contrast was also detected with a high level of accuracy (68.8%, SEp = 2.9). Lower levels of discrimination accuracy were observed for the F-P contrast (58.1%, SEp = 4.4) and the T-K contrast (56.9%, SEp = 3.1), both of which differed significantly from the reference condition.

Figure 1 illustrates the interaction between position and featural contrast that reached significance in the regression model. A robust advantage for initial over final position is apparent for the contrasts T-K (Initial: 66.9%, SEp = 4.2; Final: 46.8%, SEp = 4.5) and F-P (Initial: 67.2%, SEp = 5.9; Final: 48.3%, SEp = 6.5), with a smaller initial advantage for the S-SH contrast (Initial: 72.4%, SEp = 4.0; Final: 65.3%, SEp = 4.3). The S-T contrast was unique in showing a numerical advantage for final over initial position, but the small magnitude of the difference and large standard errors suggest that the two positions should be viewed as statistically indistinguishable (Initial: 71.0%, SEp = 5.8; Final: 76.7%, SEp = 5.5).

Figure 2 depicts the marginally significant interaction between overall accuracy and position. Among children whose experiment-wide accuracy significantly exceeded chance levels in the exact binomial test, the asymmetry between final and initial position was robust (Initial: 86.5%, SEp = 2.2; Final: 71.3%, SEp = 2.9). In the data from children whose overall response accuracy did not differ significantly from chance, there was minimal difference in discrimination accuracy across positions (Initial: 36.7%, SEp = 4.3; Final: 32.3%, SEp = 4.2).

For additional insight into the pattern seen in Figure 2, Figure 3 plots percent correct discrimination in initial versus final position for each participant, ordered by increasing accuracy on experimental trials. A dotted line represents the division of participants into the categories of chance-level versus above-chance performance, per the results of the exact binomial test. Figure 3 shows that three out of five participants in the chance-level group showed greater discrimination accuracy for contrasts in initial versus final position, while two participants showed more accurate discrimination in final position.¹⁰ Among

participants whose overall performance exceeded chance levels, 8/10 children showed at least a small numerical advantage for initial over final position.

4. Discussion

The present findings regarding differences in perceptual discrimination accuracy across featural contrasts can be compared against the results of Cutler and colleagues' (2004) study of perceptual recoverability in adults. As reported above, perceptual distance scores calculated from the fully collapsed tables reported in Cutler et al. (2004) indicated that the S-T contrast was most readily discriminated ($d = 4.72$), followed by the S-SH contrast ($d = 3.91$); the T-K contrast received a lower perceptual distance score ($d = 2.95$), followed by the F-P contrast ($d = 2.09$). If we rank these four contrasts by decreasing percent correct discrimination among child participants in the present study, virtually the same order is obtained, with the exception that the F-P contrast was discriminated with slightly greater accuracy than the T-K contrast. Thus, the present results suggest that preschool-aged children's relative sensitivity to different featural contrasts is generally congruent with that of adult listeners.

In contrast with studies of adult perception (e.g., Dorman et al., 1975; Fujimura et al., 1978; Redford & Diehl, 1999), the present results yielded no main effect of initial versus final position. However, there was a significant interaction between position and contrast. Subsequent inspection revealed that the T-K and F-P contrasts were discriminated with substantially greater accuracy in initial than final position, while the S-SH contrast showed a smaller initial advantage and the S-T contrast showed virtually no difference across positions. Variation in the strength of the positional asymmetry across different phoneme contrasts has been reported in the literature on perceptual recoverability in adults (Redford & Diehl, 1999; Wright, 2004), although there is a lack of consensus as to which contrasts are associated with greater asymmetry. Furthermore, a marginally significant interaction with the variable reflecting above-chance versus chance-level performance indicated that preschool-aged listeners with higher overall discrimination accuracy showed a reliable advantage for contrasts in initial over final position. At the individual level, at least a small numerical advantage for contrasts in initial position was evident in the responses of 11/15 children, including 8/10 participants whose overall performance exceeded chance level. In total, the present results provide no evidence that preschool-aged listeners' pattern of positional sensitivity in perception differs qualitatively from the adult pattern in which consonant contrasts are more readily recovered in initial/prevocalic position.

The finding that the advantage for initial over final position was weaker among children with chance-level performance warrants closer scrutiny in light of a maturational trend hypothesized in previous literature. The account proposed by Dinnsen & Farris-Trimble (2008, 2009) suggests that children might show more accurate perception of contrasts in final position during early stages of development, with the adult advantage for initial over final position emerging gradually over the course of maturation. It could be argued that the children with chance-level performance were the least mature overall, and that the reduced

¹⁰These two children were not among the youngest in the group; they were aged 5;0 and 5;4.

positional asymmetry exhibited by this group was a reflection of a developmentally earlier pattern of perceptual sensitivity favoring final over initial position. However, a chi-squared test indicated that the chance-level group did not differ significantly from the above-chance group with respect to the number of individuals perceiving initial contrasts with greater accuracy than final contrasts ($\chi^2 = 0.68$, $df = 1$, $p = 0.41$). Visual inspection of individual response patterns revealed that three out of five children in the chance-level group showed a moderate to large perceptual advantage for initial over final contrasts, while two children showed a moderate to large advantage for final position. These results suggest that the smaller mean difference in accuracy between final and initial position is a reflection of random fluctuations among listeners who are performing at chance level, rather than a pattern of reduced asymmetry shared across individuals in this group.

The results of the present study still do not constitute direct evidence against Dinnsen & Farris-Trimble's (2008, 2009) proposal of a child-specific pattern prominence that could yield a reversal in positional perceptual sensitivity. First, in light of the small sample size, it is possible that the present study failed to detect a significant effect of position due to insufficient experimental power. However, note that the present results revealed significant interactions and a nonsignificant trend toward greater perceptual sensitivity in initial relative to final position, the opposite of the hypothesized child-specific pattern. Second, it could be argued that the children sampled here (mean age 4;10) were old enough already to have undergone a developmental shift to an adult-like pattern of perceptual response.¹¹ However, previous studies have reported that infants show more accurate detection of phonemic contrasts in initial than final position (Altvater-Mackensen & Fikkert, 2010; Jusczyk, Goodman, & Baumann, 1999; Zamuner, 2006). A similar advantage for initial position was reported in a case study of a child aged 4;0 (McAllister Byun, 2012a). Thus, the present results can be viewed as a contribution to a growing body of empirical evidence that children across various developmental stages do not exhibit a child-specific pattern of positional sensitivity in perception.

5. Conclusion

This investigation makes a contribution to a line of research aiming to account for the phenomenon of developmental phonological patterns that lack counterparts in adult phonological typology. Within the framework of phonetically-influenced phonology, it has been argued that these child-specific phonological patterns could arise as a consequence of child-specific phonetic pressures (e.g., Inkelas & Rose, 2007; McAllister Byun, 2011). A specific claim has been advanced (Dinnsen & Farris-Trimble, 2008, 2009) that child-specific phonological patterns can be understood as the product of differences in child and adult patterns of relative prominence, with implications for perceptual recoverability of contrasts in initial versus final position. However, there has been very little empirical investigation of this hypothesis. The present study examined preschool-aged children's perceptual discrimination of four featural contexts across two positional contexts. For the contrasts and age range tested, the present study revealed no systematic evidence of divergence between

¹¹Unfortunately, no measures of lexicon size were obtained as part of the present study. Future investigations of the present topic would certainly be enhanced by collection of this information.

typical preschoolers' patterns of discrimination accuracy and the patterns of relative perceptual recoverability reported for adult listeners. Thus, the present results did not support the hypothesis of a child-specific pattern of perceptual sensitivity. This suggests that, to the extent that child speech patterns deviate from expectations based on adult patterns of perceptual recoverability, it may be more fruitful to look for an explanation in other domains. In keeping with the diversity of factors that contribute to phonological typology, these explanations could include articulatory-phonetic pressures as well as universal markedness and other influences.

Acknowledgments

The author would like to acknowledge Peter Halpin (NYU) for assistance with statistical analysis, and Susannah Levi and Lisa Davidson (NYU) for helpful comments on the manuscript. In addition, the following students assisted with the experiment: Brianna Bradley, Lindsey Syvertsen (Montclair State University); Heather Campbell, Laine Cialdella, Sydney DeBari Mary Gochuico, Deanna Kawitzky, Wendy Liang, Larissa Lum, Michelle Marron, Ashton Murray, Christopher Nightingale, and Maria Turcotte (NYU).

Funding acknowledgment

This research was supported by NIH R03DC 012883.

Appendix

Appendix A

Complete list of stimuli.

Orthography	IPA	Mean Positional Segment Frequency ¹²	Mean Biphone Frequency ¹³	Category of Contrast	Position of Contrast	Voicing of Contrast
doash-doas	/dɔʊʃ/-/dɔʊs/	0.0505	0.0023	S-SH	Final	Voiceless
fash-fass	/fæʃ/-/fæs/	0.0572	0.0034	S-SH	Final	Voiceless
heesh-heese	/hiʃ/-/his/	0.0445	0.0017	S-SH	Final	Voiceless
koish-kois	/kɔɪʃ/-/kɔɪs/	0.0405	0.0002	S-SH	Final	Voiceless
lish-liss	/liʃ/-/lis/	0.0553	0.0054	S-SH	Final	Voiceless
seeb-sheb	/seɪb/-/fɛb/	0.0546	0.0028	S-SH	Initial	Voiceless
seng-sheng	/sɛŋ/-/ʃɛŋ/	0.0536	0.0025	S-SH	Initial	Voiceless
sipe-shipe	/sɪp/-/ʃɪp/	0.0480	0.0019	S-SH	Initial	Voiceless
soadge-shoadge	/soʊdʒ/-/ʃoʊdʒ/	0.0444	0.0013	S-SH	Initial	Voiceless
soas-shoas	/soʊs/-/ʃoʊs/	0.0631	0.0028	S-SH	Initial	Voiceless
soun-shoun	/saʊn/-/ʃaʊn/	0.0630	0.0030	S-SH	Initial	Voiceless
soyt-shoyt	/soɪt/-/ʃoɪt/	0.0476	0.0000	S-SH	Initial	Voiceless
sull-shull	/sʌl/-/ʃʌl/	0.0560	0.0016	S-SH	Initial	Voiceless
vesh-vess	/vɛʃ/-/vɛs/	0.0383	0.0024	S-SH	Final	Voiceless
waysh-wase	/weɪʃ/-/weɪs/	0.0398	0.0023	S-SH	Final	Voiceless
zush-zuss	/zʌʃ/-/zʌs/	0.0330	0.0012	S-SH	Final	Voiceless
dack-gack	/dæk/-/gæk/	0.0628	0.0050	T-K	Initial	Voiced
dape-gape	/deɪp/-/geɪp/	0.0409	0.0019	T-K	Initial	Voiced
deb-geb	/deb/-/geb/	0.0466	0.0018	T-K	Initial	Voiced
duff-guff	/dʌf/-/gʌf/	0.0421	0.0029	T-K	Initial	Voiced
gud-gug	/gʌd/-/gʌg/	0.0438	0.0022	T-K	Final	Voiced
joot-jook	/dʒʊt/-/dʒʊk/	0.0377	0.0014	T-K	Final	Voiceless
koit-koik	/kɔɪt/-/kɔɪk/	0.0497	0.0001	T-K	Final	Voiceless
nowed-naug	/naʊd/-/haʊg/	0.0243	0.0005	T-K	Final	Voiced
tang-kang	/tæŋ/-/kæŋ/	0.0559	0.0061	T-K	Initial	Voiceless

Orthography	IPA	Mean Positional Segment Frequency ¹²	Mean Biphone Frequency ¹³	Category of Contrast	Position of Contrast	Voicing of Contrast
teeg-keeg	/tʰg/-/kʰg/	0.0427	0.0015	T-K	Initial	Voiceless
tep-kep	/tʰp/-/kʰp/	0.0572	0.0028	T-K	Initial	Voiceless
toab-koab	/tʰob/-/kʰob/	0.0486	0.0030	T-K	Initial	Voiceless
woad-woag	/wʰod/-/wʰog/	0.0438	0.0018	T-K	Final	Voiced
yate-yake	/jʰet/-/jʰek/	0.0397	0.0021	T-K	Final	Voiceless
yeed-yeeg	/jʰid/-/jʰig/	0.0296	0.0013	T-K	Final	Voiced
zoot-zuke	/zʰut/-/zʰuk/	0.0319	0.0010	T-K	Final	Voiceless
fazz-pazz	/fʰæz/-/pʰæz/	0.0553	0.0028	F-P	Initial	Voiceless
feb-peb	/fʰeb/-/pʰeb/	0.0513	0.0024	F-P	Initial	Voiceless
fev-feb	/fʰev/-/fʰeb/	0.0490	0.0026	F-P	Final	Voiced
koov-koob	/kʰuv/-/kʰub/	0.0442	0.0012	F-P	Final	Voiced
mife-mipe	/mʰaɪ/-/mʰaɪp/	0.0395	0.0028	F-P	Final	Voiceless
vobe-bobe	/vʰob/-/bʰob/	0.0401	0.0017	F-P	Initial	Voiced
voim-boim	/vʰum/-/bʰum/	0.0305	0.0003	F-P	Initial	Voiced
zuff-zup	/zʰaf/-/zʰap/	0.0293	0.0012	F-P	Final	Voiceless
jaus-jaut	/dʰaus/-/dʰaut/	0.0323	0.0003	S-T	Final	Voiceless
jez-jed	/dʰgez/-/dʰged/	0.0430	0.0031	S-T	Final	Voiced
seev-teev	/sʰiv/-/tʰiv/	0.0508	0.0027	S-T	Initial	Voiceless
sib-tib	/sʰib/-/tʰib/	0.0688	0.0037	S-T	Initial	Voiceless
yuzz-yud	/jʰaz/-/jʰad/	0.0367	0.0015	S-T	Final	Voiced
zeev-deev	/zʰiv/-/dʰiv/	0.0312	0.0014	S-T	Initial	Voiced
zom-dom	/zʰam/-/dam/	0.0463	0.0031	S-T	Initial	Voiced
veese-veet	/vʰis/-/vit/	0.0379	0.0018	S-T	Final	Voiceless
boin-boin	/bʰom/-/bʰom/	0.0622	0.0009	Catch (Same)	N/A	N/A
choat-choat	/tʰʊt/-/tʰʊt/	0.0460	0.0020	Catch (Same)	N/A	N/A
chod-chod	/tʰad/-/tʰad/	0.0435	0.0020	Catch (Same)	N/A	N/A
chog-chog	/tʰag/-/tʰag/	0.0320	0.0012	Catch (Same)	N/A	N/A
dabe-dabe	/dʰeb/-/dʰeb/	0.0413	0.0027	Catch (Same)	N/A	N/A

Orthography	IPA	Mean Positional Segment Frequency ¹²	Mean Biphone Frequency ¹³	Category of Contrast	Position of Contrast	Voicing of Contrast
dague-dague	/deɪg/-deɪg/	0.0403	0.0014	Catch (Same)	N/A	N/A
deek-deek	/dik/-dik/	0.0535	0.0021	Catch (Same)	N/A	N/A
div-div	/div/-div/	0.0587	0.0063	Catch (Same)	N/A	N/A
doke-doke	/dɔk/-dɔk/	0.0584	0.0028	Catch (Same)	N/A	N/A
foit-foit	/fɔt/-fɔt/	0.0418	0.0000	Catch (Same)	N/A	N/A
gaff-gaff	/gæf/-gæf/	0.0450	0.0019	Catch (Same)	N/A	N/A
gatch-gatch	/gæt/-gæt/	0.0428	0.0020	Catch (Same)	N/A	N/A
giff-giff	/gɪf/-gɪf/	0.0499	0.0022	Catch (Same)	N/A	N/A
godge-godge	/gɔdʒ/-gɔdʒ/	0.0348	0.0023	Catch (Same)	N/A	N/A
gom-gom	/gʌm/-gʌm/	0.0478	0.0038	Catch (Same)	N/A	N/A
hoague-hoague	/hɔʊg/-hɔʊg/	0.0435	0.0025	Catch (Same)	N/A	N/A
keb-keb	/kɛb/-kɛb/	0.0580	0.0017	Catch (Same)	N/A	N/A
keeve-keeve	/kiv/-kiv/	0.0481	0.0014	Catch (Same)	N/A	N/A
kiff-kiff	/kɪf/-kɪf/	0.0652	0.0038	Catch (Same)	N/A	N/A
kodge-kodge	/kɔdʒ/-kɔdʒ/	0.0502	0.0040	Catch (Same)	N/A	N/A
laug-laug	/lɔʊg/-lɔʊg/	0.0217	0.0002	Catch (Same)	N/A	N/A
lig-lig	/lɪg/-lɪg/	0.0493	0.0043	Catch (Same)	N/A	N/A
mep-mep	/mɛp/-mɛp/	0.0518	0.0034	Catch (Same)	N/A	N/A
nav-nav	/nɒ/-nɒ/	0.0423	0.0017	Catch (Same)	N/A	N/A
poin-poin	/pɔɪn/-pɔɪn/	0.0590	0.0007	Catch (Same)	N/A	N/A
rowt-rowt	/rɔʊt/-rɔʊt/	0.0418	0.0004	Catch (Same)	N/A	N/A
tague-tague	/teɪg/-teɪg/	0.0391	0.0013	Catch (Same)	N/A	N/A
teeb-teeb	/tɪb/-tɪb/	0.0393	0.0022	Catch (Same)	N/A	N/A
tiv-tiv	/tɪv/-tɪv/	0.0575	0.0026	Catch (Same)	N/A	N/A
tud-tud	/tʌd/-tʌd/	0.0562	0.0025	Catch (Same)	N/A	N/A
tull-tull	/tʌl/-tʌl/	0.0506	0.0016	Catch (Same)	N/A	N/A
vabe-vabe	/vɛb/-vɛb/	0.0241	0.0016	Catch (Same)	N/A	N/A
veese-veese	/vɪs/-vɪs/	0.0369	0.0011	Catch (Same)	N/A	N/A

Orthography	IPA	Mean Positional Segment Frequency ¹²	Mean Biphone Frequency ¹³	Category of Contrast	Position of Contrast	Voicing of Contrast
veeze-veeze	/vɪz/-vɪz/	0.0241	0.0012	Catch (Same)	N/A	N/A
vike-vike	/vaɪk/-/vaɪk/	0.0361	0.0015	Catch (Same)	N/A	N/A
voit-voit	/vɔɪt/-/vɔɪt/	0.0263	0.0001	Catch (Same)	N/A	N/A
vug-vug	/vʌg/-/vʌg/	0.0288	0.0012	Catch (Same)	N/A	N/A
wuke-wuke	/wʊk/-/wʊk/	0.0430	0.0003	Catch (Same)	N/A	N/A
yil-yil	/jɪl/-/jɪl/	0.0627	0.0062	Catch (Same)	N/A	N/A
ying-ying	/jɪŋ/-/jɪŋ/	0.0418	0.0034	Catch (Same)	N/A	N/A
yom-yom	/jam/-/jam/	0.0399	0.0021	Catch (Same)	N/A	N/A
yub-yub	/jʌb/-/jʌb/	0.0311	0.0021	Catch (Same)	N/A	N/A
zep-zep	/zɛp/-/zɛp/	0.0351	0.0009	Catch (Same)	N/A	N/A
zoid-zoid	/zɔɪd/-/zɔɪd/	0.0196	0.0000	Catch (Same)	N/A	N/A
zote-zote	/zɔʊt/-/zɔʊt/	0.0420	0.0019	Catch (Same)	N/A	N/A
zuch-zuch	/zʊtʃ/-/zʊtʃ/	0.0249	0.0008	Catch (Same)	N/A	N/A

Appendix B

Complete results of t-tests comparing properties of stimulus categories

Contrast	Mean positional segment frequency	Mean biphone frequency	Number of Neighbors
experimental vs catch	$t = -.96, df = 186, p = .34$	$t = -.41, f = 188, p = .68$	$t = -.16, df = 186, p = .88$
initial vs final	$t = 3.82, df = 80, p = .0002^{**}$	$t = 2.24, df = 94, p = .03^{*}$	$t = 1.41, df = 95, p = .16$
F-P vs T-K	$t = -.14, df = 29, p = .87$	$t = -.71, df = 44, p = .48$	$t = -1.83, df = 35, p = .07$
F-P vs S-T	$t = -.21, df = 28, p = .83$	$t = -.73, df = 27, p = .47$	$t = -1.18, df = 29, p = .25$
F-P vs S-SH	$t = -1.7, df = 42, p = .09$	$t = -.81, df = 42, p = .42$	$t = -.68, df = 33, p = .50$
S-T vs T-K	$t = .12, df = 23, p = .90$	$t = .11, df = 34, p = .91$	$t = -.34, df = 29, p = .73$
S-T vs S-SH	$t = -1.29, df = 34, p = .20$	$t = .04, df = 32, p = .97$	$t = .70, df = 27, p = .49$
TK vs S-SH	$t = -1.87, df = 53, p = .07$	$t = -.08, df = 64, p = .94$	$t = 1.34, df = 64, p = .18$

Appendix C

All fixed effects in logistic mixed model of discrimination accuracy

	β	SE	z	p
(Intercept)	-0.08	0.55	-0.14	0.89
Contrast: S-SH (reference level = S-T)	-0.71	0.51	-1.40	0.16
Contrast: T-K (reference level = S-T)	-1.49	0.52	-2.86	< 0.001
Contrast: F-P (reference level = S-T)	-1.29	0.59	-2.17	0.03
Position (reference level = initial)	-0.85	0.54	-1.57	0.12
Session	0.37	0.19	1.94	0.05
Positional segment frequency	-6.00	15.07	-0.40	0.69
Biphone frequency	187.52	112.42	1.67	0.10
Above-Chance	2.27	0.67	3.39	< 0.001
Contrast (S-SH) * Position	1.03	0.61	1.70	0.09
Contrast (T-K) * Position	1.32	0.59	2.22	0.03
Contrast (F-P) * Position	1.54	0.67	2.31	0.02
Above-Chance * Position	0.74	0.38	1.93	0.05
Above-Chance * Contrast (S-SH)	-0.17	0.59	-0.28	0.78
Above-Chance * Contrast (T-K)	-0.32	0.59	-0.54	0.59
Above-Chance * Contrast (F-P)	-0.67	0.66	-1.01	0.31

References

- Altwater-Mackensen N, Fikkert P. The acquisition of the stop-fricative contrast in perception and production. *Lingua*. 2010; 120:1898–1909.
- Aslin, RN.; Woodward, JZ.; LaMendola, NP.; Bever, TG. Models of word segmentation in fluent maternal speech to infants.. In: Morgan, JL.; Demuth, K., editors. *Signal to syntax*. Lawrence Erlbaum; Mahwah, NJ: 1996.
- Becker M, Tessier A-M. Trajectories of faithfulness in child-specific phonology. *Phonology*. 2011; 28(2):163–196.

- Beckman JN. Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology*. 1997; 14:1–46.
- Bedore LM, Leonard LB, Gandour J. The substitution of a click for sibilants: A case study. *Clinical Linguistics & Phonetics*. 1994; 8:283–293. [PubMed: 22320894]
- Bell, A.; Hooper, JB. Issues and evidence in syllabic phonology.. In: Bell, A.; Hooper, JB., editors. *Segments and syllables*. North Holland; New York: 1978. p. 9-22.
- Benkí JR. Analysis of English nonsense syllable recognition in noise. *Phonetica*. 2003; 60:129–157. [PubMed: 12853716]
- Bills, S.; Golston, C. Prosodic and linear licensing in English acquisition.. In: Carmichael, L.; Huang, C-H.; Samiiian, V., editors. *Proceedings of the Western Conference on Linguistics*. California State University, Fresno; Fresno: 2002.
- Blevins, J. *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press; Cambridge: 2004.
- Boersma, P.; Weenink, D. Praat: Doing phonetics by computer (Version 5.1.2.9). 2010. Retrieved from <http://www.praat.org/>
- Brooks PJ, MacWhinney B. Phonological priming in children's picture naming. *Journal of Child Language*. 2000; 27(2):335–366. [PubMed: 10967891]
- Chiat S. Why Mikey's right and my key's wrong: The significance of stress and word boundaries in a child's output system. *Cognition*. 1983; 14:275–300. [PubMed: 6686102]
- Coady JA, Aslin RN. Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *Journal Of Experimental Child Psychology*. 2004; 89(3):183–213. [PubMed: 15501451]
- Cooke M, García Lecumberri ML, Scharenborg O, Van Dommelen WA. Language-independent processing in speech perception: Identification of English intervocalic consonants by speakers of eight European languages. *Speech Communication*. 2010; 52:954–967.
- Cutler A, Weber A, Smits R, Cooper N. Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America*. 2004; 116:3668–3678. [PubMed: 15658717]
- Demuth K, Culbertson J, Alter J. Word-minimality, epenthesis, and coda licensing in the early acquisition of English. *Language and Speech*. 2006; 49:137–174. [PubMed: 17037120]
- Dinnsen, DA. Recalcitrant error patterns.. In: Dinnsen, DA.; Gierut, JA., editors. *Optimality theory, phonological acquisition and disorders*. Equinox Publishing Ltd; London: 2008. p. 247-276.
- Dinnsen, DA.; Farris-Trimble, AW. The prominence paradox.. In: Dinnsen, DA.; Gierut, JA., editors. *Optimality Theory, phonological acquisition and disorders*. Equinox Publishing Ltd; London: 2008. p. 277-308.
- Dinnsen, DA.; Farris-Trimble, AW. Developmental shifts in phonological strength relations.. In: Nasukawa, K.; Bäckley, P., editors. *Phonological strength*. Mouton de Gruyter; New York: 2009. p. 113-149.
- Dorman MF, Raphael LJL,AM, Repp B. Some maskinglike phenomena in speech perception. *Haskins Laboratories Status Report on Speech Research*. 1975; SR-42/43:265–276.
- Elliott LL, Hammer MA. Longitudinal changes in auditory discrimination in normal children and children with language-learning problems. *Journal of Speech and Hearing Disorders*. 1988; 53:467–474. [PubMed: 3263544]
- Fujimura O, Macchi MJ, Streeter LA. Perception of stop consonants with conflicting transitional cues: A cross-linguistic study. *Language and Speech*. 1978; 21:337–346. [PubMed: 750793]
- Gierut JA. Experimental validation of OT solutions to the comprehension–production dilemma. *Clinical Linguistics & Phonetics*. 2006; 20(7-8):485–491. [PubMed: 17056478]
- Goad, H. Consonant harmony in early grammars revisited: Domains, licensing and directionality.. Paper presented at the Journées d'Études Linguistiques (JEL) 2004: [Domain[e]s]; Nantes, France: 2004.
- Goad, H. The representation of sC clusters.. In: Oostendorp, M. v.; Ewen, C.; Hume, E.; Rice, K., editors. *The Blackwell companion to phonology*. Wiley Blackwell; Oxford, UK: 2011. p. 898-923.
- Hale M, Reiss C. *The phonological enterprise*. Oxford: Oxford University Press. 2008

- Hayes, B.; Kirchner, R.; Steriade, D., editors. Phonetically-based phonology. Cambridge University Press; Cambridge: 2004.
- Hazan V, Barrett S. The development of phonemic categorization in children aged 6- 12. *Journal of Phonetics*. 2000; 28:377–396.
- Henke, E.; Kaisse, EM.; Wright, R. Is the sonority sequencing principle an epiphenomenon?. In: Parker, S., editor. *The sonority controversy*. De Gruyter; Boston: 2012. p. 65-100.
- Hodson, B. *Hodson assessment of phonological patterns*. third edition.. Pro-Ed; Austin, Texas: 2004.
- Hoffman PR, Daniloff RG, Bengoa D, Schuckers GH. Misarticulating and normally articulating children's identification and discrimination of synthetic [r] and [w]. *Journal of Speech and Hearing Disorders*. 1985; 50:46–53. [PubMed: 3974212]
- Inkelas, S.; Rose, Y. Velar fronting revisited.. In: Beachley, P.; Brown, A.; Conlin, F., editors. *Proceedings of the 26th annual Boston University Conference on Language Development*. Cascadilla Press; Somerville, MA: 2003.
- Inkelas S, Rose Y. Positional neutralization: A case study from child language. *Language*. 2007; 83:707–736.
- Jaeger TF. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*. 2008; 59:434–446. [PubMed: 19884961]
- Juszyk PW, Goodman MB, Baumann A. Nine-month-olds' attention to sound similarities in syllables. *Journal of Memory and Language*. 1999; 40:62–82.
- Kent, RD. The biology of phonological development.. In: Ferguson, CA.; Menn, L.; Stoel-Gammon, C., editors. *Phonological Development: Models, Research, Implications*. York Press; Timonium, Maryland: 1992. p. 65-90.
- Locke JL. The inference of speech perception in the phonologically disordered child. Part II: Some clinically novel procedures; their use, some findings. *Journal of Speech and Hearing Disorders*. 1980; 45:445–468. [PubMed: 7442162]
- Luce, RD. Detection and recognition.. In: Luce, RD.; Bush, RR.; Galanter, E., editors. *Handbook of mathematical psychology*. Wiley; New York: 1963. p. 103-187.
- Macnamara, J. *Names for things: A study of child language*. Bradford Books / MIT Press; Cambridge, MA: 1982.
- Mateus, MH.; d'Andrade, E. *The Phonology of Portuguese*. Oxford University Press; Oxford: 2000.
- Mayo C, Turk A. Adult-child differences in acoustic cue weighting are influenced by segmental context: Children are not always perceptually biased toward transitions. *Journal of the Acoustical Society of America*. 2004; 115:3184–3194. [PubMed: 15237842]
- Mayo C, Scobbie JM, Hewlett N, Waters D. The influence of phonemic awareness development on acoustic cue weighting strategies in children's speech perception. *Journal of Speech, Language, and Hearing Research*. 2003; 46:1184–1196.
- McAllister Byun T. A gestural account of a child-specific neutralisation in strong position. *Phonology*. 2011; 28:371–412.
- McAllister Byun T. Bidirectional perception–production relations in phonological development: Evidence from positional neutralization. *Clinical Linguistics & Phonetics*. 2012a; 26:397–413. [PubMed: 22489733]
- McAllister Byun T. Positional velar fronting: An updated articulatory account. *Journal of Child Language*. 2012b; 39:1043–1076. [PubMed: 22225837]
- McReynolds LV, Kohn J, Williams GC. Articulatory-defective children's discrimination of their production errors. *Journal of Speech and Hearing Disorders*. 1975; 40:327–338. [PubMed: 1234946]
- Morrisette ML, Dinnsen DA, Gierut JA. Markedness and context effects in the acquisition of place features. *Canadian Journal of Linguistics*. 2003; 48:329–55.
- Nittrouer S. Learning to perceive speech: How fricative perception changes, and how it stays the same. *Journal of the Acoustical Society of America*. 2002; 112:711–719. [PubMed: 12186050]
- Nittrouer S, Studdert-Kennedy M. The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech, Language, and Hearing Research*. 1987; 30:319–329.
- Pater J. Minimal violation and phonological development. *Language Acquisition*. 1997; 6:201–253.

- Pinker, S. Language learnability and language development. Harvard University Press; Cambridge, MA: 1984.
- Redford MA, Diehl RL. The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America*. 1999; 106:1555–1565. [PubMed: 10489711]
- Rvachew S, Jamieson DG. Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders*. 1989; 54:193–208. [PubMed: 2709838]
- Shriberg LD, Kwiatkowski J, Hoffmann K. A procedure for phonetic transcription by consensus. *Journal of Speech and Hearing Research*. 1984; 27:456–465. [PubMed: 6482415]
- Slobin, DI. Cognitive prerequisites for the development of grammar.. In: Ferguson, CA.; Slobin, DI., editors. *Studies of child language development*. Holt; New York: 1973. p. 175-208.
- Steriade, D. Directional asymmetries in place assimilation: A perceptual account.. In: Hume, EV.; Johnson, K., editors. *The role of speech perception in phonology*. Academic Press; New York: 2001. p. 219-250.
- Stoel-Gammon, C. On the acquisition of velars in English.. In: Bernhardt, B.; Gilbert, J.; Ingram, D., editors. *Proceedings of the UBC International Conference on Phonological Acquisition*. Cascadia Press; Somerville, MA: 1996.
- Storkel HL, Hoover JR. An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English. *Behavior Research Methods*. 2010; 42:497–506. [PubMed: 20479181]
- Velleman SL. The role of linguistic perception in later phonological development. *Applied Psycholinguistics*. 1988; 9:221–236.
- Waldman FR, Singh S, Hayden ME. A comparison of speech-sound production and discrimination in children with functional articulation disorders. *Language and Speech*. 1978; 21:205–220. [PubMed: 732397]
- Whitehill TL, Francis AL, Ching CKY. Perception of place of articulation by children with cleft palate and posterior placement. *Journal of Speech Language and Hearing Research*. 2003; 46:451–461.
- Wright, RA. A review of perceptual cues and cue robustness.. In: Hayes, B.; Kirchner, R.; Steriade, D., editors. *Phonetically based phonology*. Cambridge University Press; Cambridge: 2004. p. 34-57.
- Ygual-Fernández A, Cervera-Mérida JF. Relación entre la percepción y la articulación en procesos fonológicos sustitutorios de niños con trastornos del lenguaje. *Revista de Neurolog*. 2013; 56:S131–S140.
- Zamuner TS. Sensitivity to word-final phonotactics in 9- to 16-month-old infants. *Infancy*. 2006; 10:77–95.

Highlights

- Some developmental phonological patterns have no counterpart in adult typology.
- Patterns might be grounded in perceptual differences between children and adults.
- This study evaluated preschool children's perception across contrasts and contexts.
- Patterns of perceptual sensitivity were broadly congruent between adults, children.

Figure 1.
Percent correct discrimination across four featural contrasts, initial versus final position.
Bars represent standard error of the proportion.

Figure 2. Percent correct discrimination in initial versus final position, partitioned by results of exact binomial test (above-chance versus chance-level performance). Bars represent standard error of the proportion.

Figure 3.

Percent correct discrimination in initial versus final position for all participants, ordered by percent correct on experimental trials. The dotted line separates participants whose overall accuracy exceeded chance level from participants performing at chance level.

Table 1

Sample stimuli from perceptual discrimination task

	Initial		Final	
	Voiced	Voiceless	Voiced	Voiceless
T-K (n = 16)	[dʌf]-[gʌf]	[toub]-[koub]	[woud]-[woug]	[zut]-[zuk]
S-SH (n = 16)	-	[ʃɛb]-[sɛb]	-	[lɪʃ]-[lɪs]
F-P (n = 8)	[vɔim]-[bɔim]	[fæz]-[pæz]	[fɛv]-[fɛb]	[maɪf]-[maɪp]
S-T (n = 8)	[zam]-[dam]	[siv]-[tiv]	[jɛz]-[jɛd]	[vis]-[vit]
Catch trials (n = 46)	[næv]-[næv],[tʌd]-[tʌd],[jɪl]-[jɪl]			

Table 2

Summary of individual results from perceptual discrimination task. Rows are ordered based on overall percent correct, from lowest to highest.

Participant	Age	Hit Rate	False Alarm Rate	Overall Percent Correct	Exact binomial p
F6	5;0	46.9	51.1	47.9	0.76
F1	3;11	19.6	17.3	50.0	1.00
M1	3;6	22.0	13.0	53.1	0.61
M6	5;5	39.6	27.1	56.3	0.26
F8	5;4	46.9	29.8	58.3	0.13
M3	4;9	58.3	35.4	61.5	0.03
M5	5;3	50.0	4.2	72.9	<.001
M4	5;3	75.5	17.0	79.2	<.001
F5	4;11	85.4	22.9	81.3	<.001
F7	5;4	85.7	21.3	82.3	<.001
F3	4;7	77.6	6.4	85.4	<.001
F9	5;7	84.0	10.9	86.5	<.001
M2	4;7	84.0	8.7	87.5	<.001
F2	4;7	92.0	13.0	89.6	<.001
F4	4;9	95.8	8.3	93.8	<.001