In: A. Biller, E. Chung & A. Kimball (eds.) Proceedings of the 36th Annual Boston University Conference on Language Development. Volume 1. Somerville, MA: Cascadilla Press, 126-138.

BUCLD 36 Proceedings
To be published in 2012 by Cascadilla Press
Rights forms signed by all authors

Generalizing without Encoding Specifics: Infants Infer Phonotactic Patterns on Sound Classes

Alejandrina Cristia and Sharon Peperkamp

1. Introduction

In the study of language acquisition, a key question involves the representations that infants use while processing speech and learning phonology. Precise representations for specific sounds are crucial when the infant starts putting together a lexicon, since minimal distinctions can cue differences in meaning. In the process of language acquisition, infants must also learn constraints on the position and sequencing of sounds, that is, the *phonotactics*. Many - if not most - phonological patterns concern sets of sounds that share phonetic and phonological characteristics, rather than individual sounds or groups of completely unrelated sounds (Mielke, 2008). For instance, in German and Russian, all final stops and fricatives are voiceless; but in no language is it the case that only some stops and some fricatives are voiceless word-finally, whereas the others are voiced. Based on such generalizations, phonological theory has often assumed that learners can represent phonotactic patterns directly in terms of the natural class, for example through the phonological features that are shared between all the members of the class and no non-member (Chomsky and Halle, 1968; Hall, 2001). In psycholinguistic terms, representing phonotactic patterns at the level of the sound class may be particularly advantageous, as this allows both a more parsimonious description and the generalization to unobserved cases. In this paper, we assess whether infants' *early* representations of phonotactic patterns are based on sound classes, or whether class-based representations are derived from sound-based constraints.

A growing literature using artificial grammar paradigms strongly suggests that young infants *can* encode patterns on sound classes (a recent summary in Cristia et al., 2011a). In these studies, infants are first exposed to a phonological pattern in a subset of a natural class, and subsequently tested with the pattern instantiated

* This work was carried out at the Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS, ENS-DEC, CNRS, thanks to the financial support of Ecole de Neurosciences de Paris and Fondation Fyssen to AC; ANR 2010-BLAN-1901 and Fondation de France to SP.

in untrained sounds which either belong to that natural class or not. The experimentally controlled exposure allows us to ensure that, if infants are sensitive to the pattern even in this untrained context, then it must be because they have somehow internally generalized the exposed pattern, since this evidence has been withheld from infants' experience. To take a specific example, 7-month-old infants who had heard about a set of nonwords, all of which began with a subset of non-continuant consonants (/b k m n/), showed a significant preference at test when presented with novel non-continuant consonants (such as /t/) as opposed to novel continuant consonants (e.g., /s/; Cristia and Seidl, 2008).

Recent work also suggests that infants can encode patterns on individual onsets that do not form a natural class. Chambers et al. (2011) exposed 10-month-olds to a set of nonwords, all of which began with one of /b k n f/. At test, infants looked significantly longer when they were presented with nonwords having novel onsets (/p g m s/) than with novel nonwords with the familiar onsets. Since the familiarization and test onset sets did not differ on any single feature, infants could not succeed by learning a class-wide constraint, but must have encoded the phonotactic pattern on the individual onsets in enough detail to effect the discrimination at test.

In view of this evidence, one must ask: How do infants arrive at class-wide constraints? There are two hypotheses, which we will call *sequential* and *parallel* respectively. According to the *sequential* hypothesis, infants encode the phonotactics of each onset separately, and subsequently generate a more abstract, class-based representation. To take a specific example, in the Cristia and Seidl (2008) study mentioned above, 7-month-olds would have learned 4 constraints (on /b/, /k/, /m/, and /n/), and then worked out that these 4 sounds are all non-continuant. Alternatively, infants may perform statistics based on features (or classes) and segments in *parallel*. The latter hypothesis makes an interesting prediction: In the case that all onsets provide evidence on the same class, the class-wide pattern should be learned more quickly than the segment-specific ones. To continue with the Cristia and Seidl (2008) example, infants heard 14-15 non-words of each onset /b k m n/, but all 57 training non-words began with a non-continuant sound. Therefore, there were 4 times as many tokens evidencing the class-wide pattern than the segment-specific ones, which should have made the class-wide pattern evident even when the segment-specific patterns are not.

One study appears to speak to this question. Using a similar (but not identical) paradigm to that detailed above, Saffran and Thiessen (2003) studied 9-month-olds' phonotactic processing. In that paper, one group of infants was able to learn that all items began with one of /p t k/ (or /b d g/), whereas another group, exposed to the same number of tokens and for just as long, was not able to learn that all items began with one of /p d k/ (or /b t g/).[1] This may indicate that, with a similar amount of exposure, infants can learn a class-wide constraint but not a set of

[1] In fact, Saffran and Thiessen (2003) included an intermediate segmentation phase, which may have rendered the task more difficult. See Chambers et al. (2011) for further discussion.

segment-specific patterns, as predicted by the parallel hypothesis. However, there is an alternative interpretation to this finding, namely that learners have a cognitive bias against arbitrary classes (Pycha et al., 2003; but see Cristia et al., 2011b).

On the whole, then, it is still unclear whether infants' representations of phonological patterns are derived from, and dependent on, previous encoding of the patterns in terms of the individual segments. We carried out two experiments to directly address this question.

2. Experiment 1: Trained versus Legal

In the first experiment, infants were exposed to nonwords whose onsets were restricted to 3 obstruents sharing voicing (but not place or manner).[2] If infants in the current study spontaneously encoded the 3 separate phonotactic patterns (one for each consonant), they should also exhibit a significant preference at test, even if the competitors belong to the same natural class encompassing those 3 onsets. Alternatively, if infants directly encode the phonotactic pattern in terms of classes, the distinction between familiar and novel within-class items becomes challenging.

2.1. Participants

Thirty-six 6-month-olds (M= 6,15; range 5,28-6,29; 24 females) were included in the final analyses. An additional 10 infants were excluded for: unreliable online coding (5)[3]; not completing the experiment due to fussing (2); equipment error (1); experimenter error (1); having a health problem at the time of test (1). All infants were primarily exposed to Parisian French.

[2]This may be described as a first order constraint, where a segment or class is constrained to a given syllabic position regardless of what other sounds are present in the same nonword. An example of such a constraint is the voicing patterns found in German and Russian (word-final obstruents must be voiceless, regardless of which sounds or classes precede that final obstruent). Some previous studies looking at first order constraints have used the constrained sounds in other syllable positions (e.g., in Cristia and Seidl, 2008 onsets could only be non-continuant sounds, but codas could be continuant or non-continuant consonants), whereas others - like ours - do not use the constrained sounds in any other position (e.g., in Chambers et al., 2011 /b k n f/ only occur as onsets and no other onset is possible). It is unclear whether this experimental decision impacts infants' performance, since (a) evidence that infants attend to non-initial phonotactics is scarce (perhaps limited to Seidl et al., 2009, where the exposure was twice as long as in the present study; see e.g., Jusczyk et al., 1999 and Zamuner, 2006 for further discussion of infants' inattention to word offsets); and (b) evidence that infants map word-initial and word-final sounds onto the same category is non-existent. In the latter case, infants could also succeed at test if they responded to the base frequency of the sounds/classes involved.

[3]A research assistant was being trained, and her online coding did not always follow the infant gaze. Coding was considered unreliable if a trial was triggered before the child fixated on the screen, and/or if the trial stopped before the child looked away.

2.2. Stimuli

Items were of the form $CV_1N_1V_2(N_2)$, where C, the onset, was an obstruent (/p,t,k,f,s,ʃ,b,d,g,v,z,ʒ/), V_1 in /a,e,i,o,u/, N_1 in /m,n,l/, V_2 in /a,i,u/, and an occasional N_2 in /m,l/. A final coda was added when the $CVNV$ combination was a real French word; this coda was /l/ unless C_1 was also /l/, in which case the coda was /m/. In addition, codas were added to balance the proportion of codas across sets of stimuli heard by individual infants. None of the items was a real word in French. There were 324 items overall, 27 for each of the 12 onsets; 40% of the items were consonant-final. Two thirds of the items were used for the familiarization phase, and a third were reserved for the test phase; each of these sets contained 18 items per onset.[4]

2.3. Procedure

Infants were randomly assigned to one of 12 counterbalancing groups, as shown in Figure 1. Each group was familiarized to items with three different onsets that shared their voicing value; the onsets consisted either of two stops and one fricative, or of one stop and two fricatives. Exposure consisted of 54 different non-words, 18 per onset. During the test phase, infants listened to 6 trials with 9 different items each. A third of the items was reserved for test for two reasons. First, in this way, all infants were faced with novel items at test. Additionally, the same set of stimuli was reserved for all infants so that any difference across conditions would only be attributable to the initial training. Within individual test trials, all items had the same onset. In half of the trials ($familiar$ trials), the onsets were the same as those during familiarization (but the items were novel); in the other half ($legal$ trials), the onsets were novel but they shared the voicing value of the familiarization onsets. Across the counterbalancing groups, each of the 12 onsets served as familiar and legal the same number of times, and there were no differences in diphone frequency or in proportion of consonant-final items in the sets used for familiarization and testing.

To take two specific examples, infants in Counterbalancing Group 1 heard 18 different non-words beginning with /b/, 18 with /d/, and 18 with /ʒ/. At test, they heard 6 trials, one for each of /b d ʒ/ (the *familiar* onsets) and of /g v z/ (the *legal* onsets). Infants in Counterbalancing Group 2 heard 18 different non-words of each of /g v z/ during familiarization, and were tested with exactly the same trials as infants in Counterbalancing Group 1, but the trials mapped onto the opposite categories (/b d ʒ/ were *legal* and /g v z/ *familiar*).

The infant sat on a caregiver's lap in the center of a small room, in front of a screen on which images were shown; the sound was produced from speakers located behind the screen. During familiarization, the stimulus list was played twice in a pseudo-random order, and regardless of whether the infant focused on

[4] Stimuli and data are available for download. Please visit sites.google.com/site/acrsta for further details.

Group		voiceless						voiced							voiceless						voiced					
		stop			fric			stop			fric				stop			fric			stop			fric	
		p	t	k	f	s	ʃ	b	d	g	v	z	ʒ		p	t	k	f	s	ʃ	b	d	g	v	z	ʒ
1	Familiarization							F	F				F	Test							F	F	L	L	L	F
2										F	F	F									L	L	F	F	F	L
3									F	F	F										L	F	F	F	L	L
4								F				F	F								F	L	L	L	F	F
5									F		F		F								L	F	L	F	L	F
6								F		F		F									F	L	F	L	F	L
7		F	F				F								F	F	L	L	L	F						
8				F	F	F									L	L	F	F	F	L						
9			F	F	F										L	F	F	F	L	L						
10		F				F	F								F	L	L	L	F	F						
11			F		F		F								L	F	L	F	L	F						
12		F		F		F									F	L	F	L	F	L						

Figure 1: Attribution of consonants to familiarization and test trial types across counterbalancing groups in Experiment 1. See main text for details.

the screen. During this phase, which lasted about 2 minutes, visuals were generated with the iTunes visualizer. The structure of the 6 test trials was as follows. First, a moving attention getter was shown on the screen until the child focused on it. When the child looked, a static image (a bull's eye for 24 of the children; a checkerboard for the other 12) was projected onto the screen and the auditory stimuli began. The audiovisual presentation continued until one of the following conditions was met: (a) the infant looked away for more than 1 second; (b) the auditory stimuli had been presented 3 times. The experimenter monitored the child's looks through a closed video circuit; videos were digitized at 25 frames per second and coded offline. The dependent measure was *looking time*, the amount of time the infant fixated on the screen while hearing the test stimuli; this was averaged across the 3 trials of each type.

2.4. Results

Average looking times to familiar trials was 5.06 s (SD=3.01), to legal trials 5.17 s (SD=3.24); these averages are shown on Figure 2 (left). This difference was not reliable, $t(35)=.2$, $p=.85$; Cohen's $d=.05$. Only 17 of the 36 infants looked longer during legal trials than during familiar trials, which was not significant according to a binomial test, $p > .05$.

2.5. Discussion

After being familiarized with 54 non-words, which began with 3 obstruents of the same voicing, infants in this experiment did not show a reliable preference for legal trials, where non-words began with 3 new, untrained obstruents, over familiar trials, where novel non-words began with the 3 familiar onsets. This is exactly the pattern of results predicted by the parallel hypothesis, as detailed in the introduction.

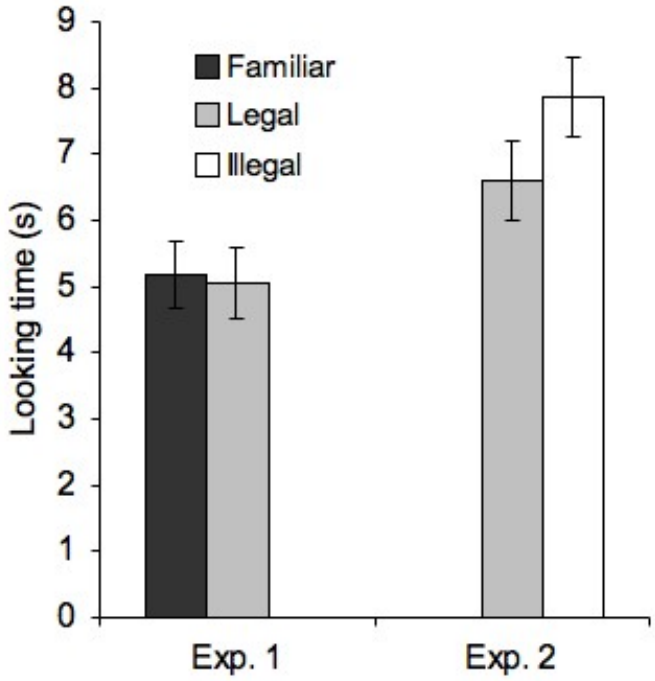

Figure 2: Looking times by trial type in Experiment 1 and Experiment 2.

However, it is also possible that infants failed to show a preference here for extraneous reasons. A first concern is whether learning statistics on 3 consonants may have proved too hard. For example, 10-month-olds in Chambers et al. (2011) learned statistics on 4 onsets after hearing 80 tokens (16 different non-words, repeated 5 times). Infants in the present study heard more tokens (108: 54 different non-words x 2 repetitions), but they were considerably younger. Nonetheless, age does not appear to be a likely candidate to explain the null effect, since 4-month-olds in both Seidl et al. (2009) and Cristia et al. (2011b) successfully learned a phonotactic constraint (and even generalized it) after a mere 57 tokens.

One could still argue that those 4-month-old studies are not relevant, in view of some evidence that the linguistic development of infants learning Parisian French is different. Indeed, one study has found that French learners in Canada (Polka and Sundara, in press) are able to perform word segmentation tasks at a similar age as American English learning infants, but several months *younger* than French learners in Paris, France (Nazzi et al., 2006). Therefore, even though many studies have shown that English-learning infants in the American Midwest and French-learning infants in Quebec can learn novel sound patterns after exposures of a

similar length and composition to those used here, Parisian infants could be delayed compared to their North American peers.

3. Experiment 2: Legal versus Illegal

We carried out a second experiment to address the alternative interpretation according to which infants failed in Experiment 1 due to extraneous reasons. Infants in Experiment 2 were exposed to the exact same training phases as those in Experiment 1. At test, however, they were presented with only novel onsets. Some of the onsets belonged to the familiarized $class$ (legal onsets, exactly as in Experiment 1), and others did not (these were illegal onsets). If Parisian infants are delayed in phonotactic learning, then no preference should arise in this case either. Contrastingly, if the null result in Experiment 1 followed from infants' encoding patterns at a class-wide level, then a significant preference should ensue.

3.1. Participants

Thirty-six 6-month-olds (M= 6,17; range 5,26-7,8; 20 females) were included in the final analyses. An additional 15 infants were excluded for: unreliable online coding (6; see footnote 3); experimenter error (3); having a health problem at the time of test (5); sibling interference (1). The language background of these infants was comparable to that of infants in Experiment 1.

3.2. Stimuli

The stimuli were the same as those used in Experiment 1.

3.3. Procedure

The procedure was the same as that in Experiment 1, except for the test stimuli as follows. During the test phase, half of the trials again contained onsets that shared the voicing value of those used in familiarization but were otherwise different ($legal$); the other half, however, contained the voicing-changed counterparts of these onsets ($illegal$) (see Figure 3). For instance, infants in Counterbalancing Group 1 heard the exact same training as infants in Counterbalancing Group 1 of Experiment 1 (18 different non-words beginning with /b/, 18 with /d/, and 18 with /ʒ/, repeated twice). At test, they heard 6 trials, 3 of which were exactly the same that infants in Group 1 of Experiment 1 had heard, namely /g v z/ (the *legal* trials); the other 3 were different from what their peers had heard (namely /k f s/).

3.4. Results

Average looking times to legal trials was 6.61 s (SD=3.56), to illegal trials 7.87 s (SD=3.63). This difference was significant, $t(35)=2.2$, $p=.03$; Cohen's

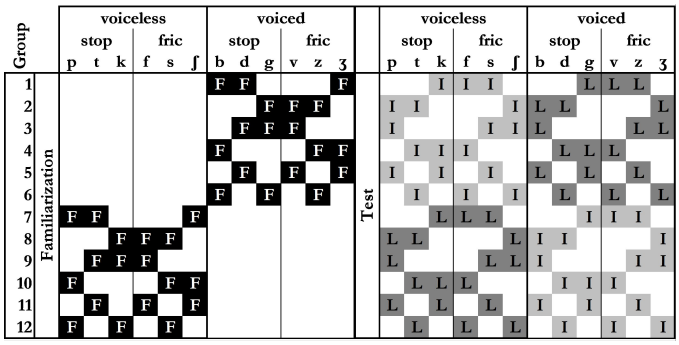

Figure 3: Attribution of consonants to familiarization and test trial types across counterbalancing groups in Experiment 2. See main text for details

d=.48. Twenty-four of the 36 infants looked longer during illegal trials than during legal trials, which was significant according to a binomial test, p=.02.

3.5. Discussion

After hearing the same familiarization as infants in infants in Experiment 1, those in Experiment 2 showed a significant preference for illegal over legal trials. This rules out the possibility that the failure in Experiment 1 was due to extraneous reasons, such as language development of the infants tested or eventual peculiarities of the stimuli used. Thus, there is clear evidence that infants exposed to this material formed representations that allowed them to distinguish between legal and illegal onsets, but not two types of legal onsets (familiar versus legal). This is consistent with the parallel hypothesis laid out above: When a new set of phonotactics is most easily described as a class-based constraint, infants do not encode the specific segments but *directly* the class-wide pattern.

We do not mean to say that infants encode the phonotactics they hear only in terms of classes. As we mentioned above, Chambers et al. (2011) clearly demonstrate that 10.5-month-old infants can learn a simple phonotactic pattern on an arbitrary set of sounds, replicating their earlier studies on 16.5-month-olds (Chambers et al., 2003). To learn such patterns, infants must have encoded the phonotactics of each sound in the arbitrary set separately. What remains to be explained is why infants encoded each sound separately in Chambers et al. (2003, 2011) but not in our Experiment 1. More in general, what are the conditions that lead infants to encode primarily class-wide constraints?

One possibility relates to the fact that stimuli were designed in very different ways across the two strands of studies. In the present study, the 'bodies' of the nonwords varied a great deal while all initial onsets were fairly similar, and items were only presented twice. In a Chambers-type design, by contrast, the items are few and not very variable, and they are presented many times. Other work

suggests that more variability leads infants to focus on the more constant aspects in the speech sounds they hear: Twelve-month-olds in Gomez and Maye (2005) were able to learn a non-adjacent co-occurrence pattern only if the element separating the co-occurring elements was highly variable. Therefore, it is possible that the conjunction of item repetition and low variability led infants in Chambers et al. (2003, 2011) to encode segment-specific constraints, whereas low item repetition and high variability allowed class-based learning in the present study. We sought further evidence on this question by assessing correlations between experimental design and effect sizes in previously published work on infant phonotactic learning, which are tabulated in Table 1. We made 2 predictions to test in these previous data: (1) that the number of types would correlate more closely with generalization performance (legal versus illegal, or L–I) than with learning performance (familiar versus illegal, or F–I); and (2) that the number of repetitions would correlate more closely with learning performance than with generalization performance. Neither prediction was born out (1: correlation between Types and F–I r =.92, df=8, p <.01; Types and L–I r =-.33, df=10, n/s; 2: Repetitions and F–I r =-.73, p <.01; Repetitions and L–I r =.2, n/s). Hence, there was no supporting evidence for our predictions based on the hypothesis that differences in variability in the exposure stimuli may explain when infants encode exposure patterns in terms of individual sounds or in terms of classes. Naturally, more direct evidence may be sought through direct manipulation in follow-up studies, by varying the composition of the training stimuli.

A second explanation relates to infants' ages: It is possible that younger infants are biased towards class-based encoding, whereas older learners readily encode segment-based constraints. The data to answer this question is scarce, since most studies assessing generalization have been carried out with young infants, and most studies assessing segment-specific learning were carried out with older infants. The only exceptions are 2 studies with 14-month-olds on generalization (Cristia, 2006; Cristia et al., 2011a), both of which yielded null results, and 3 studies with 10.5-month-olds on segment-specific learning, two of which yielded non-significant effects. Insofar as null studies can support a hypothesis, it does appear that older infants find it harder to encode class-wide patterns, whereas younger infants struggle with segment-based patterns.

Going even further in terms of development, it is clear from previous research that adults do not readily and spontaneously generalize newly experienced phonotactics to untrained within-class sounds. For instance, Peperkamp et al. (2006), Peperkamp and Dupoux (2007), and Finley (in press) found no evidence of generalization, whereas Finley and Badecker (2009) documented generalization in 2 out of 3 sound patterns tested. Furthermore, in recent work, we assessed adults' generalization with the same stimuli and a similar design to that used here (Cristia et al., submitted). In this study, French adults did not encode the experienced patterns on the basis of sound classes, and their generalization was based on individual sound distances. That is, after training with /p,t,k,f,ʃ/, adults treated /s/ and

Table 1: Stimuli design and effect sizes in studies on infants' phonotactic learning. See the bibliography for the full references; the numbers and letters at the end of each reference uniquely identify the experiment within the original article. Reps=number of repetitions of each item; F=Familiar; L=Legal; I=Illegal. The last 3 columns show the t or F value reported in the original articles. Notes: [1] The 'minus' codes that the preference is in the opposite direction to what has been recorded for that age group and pattern complexity; see the original article for further details. [2] Since Seidl and Buckley (2005) mixed familiar and legal items in the same trials, their result is scored for both the familiar-illegal and the legal-illegal comparisons. [3] Seidl et al. (2009) tested 4-month-olds on both familiar-illegal and legal-illegal comparisons.

			First order patterns					
Reference	Age	N	Repetitions	Types	Tokens	F–I	F–L	L–I
Chambers et al. (2003) 1	16.5	8	6	25	150	2.65		
Chambers et al. (2003) 1	16.5	8	5	25	125	2.37		
Chambers et al. (2011) 1	10.5	16	5	16	80	2.67		
Cristia (2006) 3	14	22	1	0	0			0.12
Cristia and Seidl (2008) 1S	7	12	1	57	57			3.06
Cristia and Seidl (2008) 1F	7	12	1	57	57			-1.39[1]
Cristia and Seidl (2008) 2	7	24	1	26	26			4.51
Cristia et al. (2011a)	8	40	1	57	57			6.57
Cristia et al. (2011a)	14	36	1	57	57			0.09
Cristia et al. (2011b) S	4	24	1	57	57			2.03
Cristia et al. (2011b) F	4	24	1	57	57			2.82
This article 1	6	36	2	54	108		0.2	
This article 2	6	36	2	54	108			2.2
			Second order patterns					
Chambers et al. (2011) 2	16.5	8	5	25	125	2.92		
Chambers et al. (2011) 2	16.5	8	5	25	125	2.49		
Chambers et al. (2011) 2	10.5	16	5	16	80	0.91		
Chambers et al. (2011) 2	10.5	32	5	16	80	1.87		
Seidl and Buckley (2005) 1	9	24	1	48	48	6.76		6.76[2]
Seidl and Buckley (2005) 2	9	22	1	48	48	8.14		8.14[2]
Seidl et al. (2009) 1	11	18	5	36	180			10.44
Seidl et al. (2009) 2	11	36	5	36	180			1.82
Seidl et al. (2009) 3	4	36	5	36	180	7.48		4.96[3]

/ʒ/ similarly; notice that /s/ belonged to the training set, but both /s/ and /ʒ/ were 1 feature away from at least 1 of the onsets in the training set. Thus, development and further linguistic experience (additional exposure to a single language's phonotactics, as well as babbling, talking, lexical pressure, and possibly even orthography) may have a profound effect on the perceptual units that serve to represent sound patterns. Future work with older infants, children, and adults varying in these experiences may shed light on how they influence learners' encoding of newly learned phonotactics.

Regardless of changes that occur with further development/experience, our results suggest that infants who are in the process of learning their language's phonotactics may exhibit specific biases in their encoding of those sound patterns. Specifically, based on the findings above, it is expected that young infants will first pick up on the class-based phonotactics of their ambient language, and only later on the segment-based constraints. Let us take a specific example from two English restrictions. One concerns /ŋ/, which only occurs at the end of syllables. The other is assimilation for place of articulation when nasals co-occur with stops. These two phonotactic patterns occur with similar surface frequency (3.67% and 3.47% of the time in the CELEX corpus; Baayen et al., 1995). However, the two patterns differ in the complexity of the evidence offered to the infants. The /ŋ/ pattern is segment-specific, because the other nasals, and the other dorsals/velars, do occur syllable-initially, providing counter-evidence at the level of the class. In contrast, all class-based evidence aligns for the nasal place assimilation phonotactic pattern. The prediction made from current results is that infants should find it harder to learn the /ŋ/ pattern. We hope future research may investigate this hypothesis (although it may be difficult to gather the relevant evidence; see footnote 2).

4. Conclusion

We have documented that young infants encode phonotactics in terms of classes without encoding the specific sounds that instantiated those classes. More generally, it is possible that this bias facilitates phonotactic learning, given that most sound patterns concern classes of sounds. In a review of 500 languages, Mielke (2008) found that at least 75% of sound patterns are parsimoniously captured through sound classes. Naturally, this cross-linguistic tendency for sounds to have similar phonological behavior could be described both as an effect of a bias to learn class-based patterns (which should be effective in early infancy; see also Cristia et al., 2011a) and as a situation in which this bias is most useful since it fits the characteristics of the input. It is difficult to tease these two possibilities apart, but for the time being we conclude that young infants show a bias to represent class-based patterns, since they learn general phonotactic patterns without memorizing the specific members of the sound class.

References

Baayen, Harald R., Richard Piepenbrock, and Leon Gulikers. 1995. The CELEX lexical database [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Chambers, Kyle, Kristine H. Onishi, and Cynthia Fisher. 2003. Infants learn phonotactic regularities from brief auditory experience. *Cognition* 87:B69–B77.

Chambers, Kyle E., Kristine H. Onishi, and Cynthia Fisher. 2011. Representations for phonotactic learning in infancy. *Language Learning and Development* 7:287–308.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of english*. New York: Harper and Row.

Cristia, Alejandrina. 2006. Speech sound categories in language acquisition and learning. Unpublished Masters of Arts Thesis, Purdue University.

Cristia, Alejandrina, Jeff Mielke, Robert Daland, and Sharon Peperkamp. submitted. Constrained generalization of implicitly learned sound patterns .

Cristia, Alejandrina, and Amanda Seidl. 2008. Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development* 5:203–227.

Cristia, Alejandrina, Amanda Seidl, and Alexander Francis. 2011a. Phonological features in infancy. In *Where do phonological contrasts come from? Cognitive, physical and developmental bases of phonological features*, ed. G. Nick Clements and Rachid Ridouane, 303–326. John Benjamins.

Cristia, Alejandrina, Amanda Seidl, and LouAnn Gerken. 2011b. Young infants learn sound patterns involving unnatural sound classes. *University of Pennsylvania Working Papers in Linguistics* 17:9.

Finley, Sara. in press. Generalization to novel consonants in artificial grammar learning. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* .

Finley, Sara, and William Badecker. 2009. Artificial language learning and feature-based generalization. *Journal of Memory and Language* 61:423–437.

Gomez, Rebecca, and Jessica Maye. 2005. The developmental trajectory of nonadjacent dependency learning. *Infancy* 7:183–206.

Hall, T. Alan. 2001. Introduction: Phonological representations and phonetic implementation of distinctive features. In *Distinctive feature theory*, ed. T. Alan Hall, 1–40. Berlin and New York: Mouton de Gruyter.

Jusczyk, Peter W., Mara B. Goodman, and Angela Baumann. 1999. Nine-month-olds' attention to sound similarities in syllables. *Journal of Memory and Language* 40:62 – 82.

Mielke, Jeff. 2008. *The emergence of distinctive features*. New York: Oxford University Press.

Nazzi, Thierry, Galina Iakimova, Josiane Bertoncini, Severine Frédonie, and Carmela Alcantara. 2006. Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language* 54:283–299.

Peperkamp, Sharon, and Emanuel Dupoux. 2007. Learning the mapping from surface to underlying representations in an artificial language. In *Laboratory Phonology IX*, ed. Jennifer Cole and Jose-Ignacio Hualde, 315–338. Berlin: Mouton de Gruyter.

Peperkamp, Sharon, Katrin Skoruppa, and Emanuel Dupoux. 2006. The role of phonetic naturalness in phonological rule acquisition. In *Proceedings of the 30th Annual Boston University Conference on Language Development*, ed. D. Bamman, T. Magnitskaia, and C. Zaller, 464–475. Cascadilla Press.

Polka, Linda, and Megha Sundara. in press. Word segmentation in monolingual infants acquiring Canadian English and Canadian French: Native language, cross-dialect, and cross-language comparisons. *Infancy* .

Pycha, Anne, Pawel Nowak, Eurie Shin, and Ryan Shosted. 2003. Phonological rule-learning and its implications for a theory of vowel harmony. In *WCCFL 22 Proceedings*, ed. G. Garding and M. Tsujimura, 101–114. Cascadilla Press.

Saffran, Jennifer, and Eric D. Thiessen. 2003. Pattern induction by infant language learners.

Developmental Psychology 39:484–494.

Seidl, Amanda, and Eugene Buckley. 2005. On the learning of arbitrary phonological rules. *Language Learning and Development* 1:289–316.

Seidl, Amanda, Alejandrina Cristia, Kristine H. Onishi, and Amelie Bernard. 2009. Allophonic and phonemic contrasts in infants learning of sound patterns. *Language Learning and Development* 5:191–202.

Zamuner, Tania S. 2006. Sensitivity to word-final phonotactics in 9- to 16-month-old infants. *Infancy* 10:77–95.