Journal of Communication Disorders 68 (2017) 70–80

Contents lists available at ScienceDirect

Journal of Communication Disorders

journal homepage: www.elsevier.com/locate/jcomdis

# Perceptual statistical learning over one week in child speech production

Peter T. Richtsmeier[a,*], Lisa Goffman[b]

[a] Department of Communication Sciences and Disorders, Oklahoma State University, United States
[b] Callier Center for Communication Disorders, University of Texas at Dallas, United States

ARTICLE INFO

*Keywords:*
Statistical learning
Speech sounds
Phonological development
Child speech
Speech production

ABSTRACT

What cognitive mechanisms account for the trajectory of speech sound development, in particular, gradually increasing accuracy during childhood? An intriguing potential contributor is statistical learning, a type of learning that has been studied frequently in infant perception but less often in child speech production. To assess the relevance of statistical learning to developing speech accuracy, we carried out a statistical learning experiment with four- and five-year-olds in which statistical learning was examined over one week. Children were familiarized with and tested on word-medial consonant sequences in novel words. There was only modest evidence for statistical learning, primarily in the first few productions of the first session. This initial learning effect nevertheless aligns with previous statistical learning research. Furthermore, the overall learning effect was similar to an estimate of weekly accuracy growth based on normative studies. The results implicate other important factors in speech sound development, particularly learning via production.

## 1. Introduction

This paper addresses expressive speech sound development—or how children come to produce the sound patterns of their native language—within the framework of statistical learning. We review a large and growing literature on frequency effects in child speech. These studies, as well as closely related corpus work, suggest that statistical learning could be a prominent factor in expressive speech sound development. However, only a handful of studies have applied the statistical learning framework to child speech accuracy. Thus, there is an empirical gap separating our understanding of statistical learning on the one hand and child speech development on the other hand. We begin with a basic and longstanding finding from the child speech literature: Children gradually become more accurate in their productions of speech sounds. The purpose of this study is to explore the relevance of statistical learning to that finding.

## 2. Statistical learning in language and speech development

It may be helpful to establish a working definition of statistical learning. Romberg and Saffran (2010) refer to statistical learning as the “discovery of patterns in the input” (p. 906). They apply the term widely, for example, to the operant conditioning of behaviorism and to learning algorithms implemented on a computer. This broad application of the term reflects statistical learning’s apparent relevance to all linguistic patterns that are inherently statistical, or frequency-based.

* Corresponding author at: Communication Sciences and Disorders, 042 Murray Hall, Stillwater, OK 74078, United States.
*E-mail address:* prichtsmeier@yahoo.com (P.T. Richtsmeier).

http://dx.doi.org/10.1016/j.jcomdis.2017.06.004
Received 30 September 2016; Received in revised form 9 June 2017; Accepted 13 June 2017
Available online 17 June 2017
0021-9924/ 

Statistical learning research over the last 20 years has primarily focused on infant perceptual development. This research has adhered to a basic procedure: Infants are familiarized with an abbreviated or artificial language containing statistical regularities. Following the familiarization, infants are exposed to test items that conform to or do not conform to those regularities. Based on differences in looking times to the conforming and nonconforming items, researchers infer what infants have learned. For example, Saffran, Aslin and Newport (1996) examined statistical learning of nonsense words like *dutaba* and *pidabu* embedded in a larger, continuous stream of syllables like *dutabatutibupidabupatubibupadababupudutaba*. Following exposure to the syllable stream, infants looked longer when hearing part-words like *tabatu* compared to words like *dutaba*, indicating that infants were sensitive to which syllables went together to form words.

Based on a large body of infant research, statistical learning appears to be relevant to many aspects of language acquisition (for in-depth reviews see Gómez & Gerken, 2000; Romberg & Saffran, 2010). Furthermore, infant statistical learning persists over time (Gómez, 2011) and operates in cognitive domains besides language, for example, in visual pattern learning (Fiser & Aslin, 2002).

Returning to expressive speech sound development, statistical learning could be defined as the discovery of basic speech patterns, including the production of sounds and sound sequences. There is reason to believe that the type of statistical learning studied in infants is relevant to these patterns, that is, to expressive speech sound development. In fact, a considerable literature has documented how child speech accuracy is closely tied to the statistics of the child's native language. For example, Zamuner, Gerken and Hammond (2004) reported that two-year-olds produce a coda consonant like /-n/ more accurately if that consonant is in a high probability environment, as in /nin/, compared to a relatively low probability environment, as in /von/. Munson (2001) controlled the frequency of word-medial consonant sequences in nonwords like /fæstət/ and /fæfpət/. He found that both three-year-olds and eight-year-olds were faster, more accurate, and less variable when producing high English frequency sequences like /st/ compared to low frequency sequences like /fp/. Storkel (2001) also found that children were more accurate when producing novel words composed of high probability sequences, like /hʌp/, compared to words composed of lower probability sequences, like /gim/. Furthermore, children in Storkel's study were better able to match the high probability wordforms to semantic referents and to later recall that semantic knowledge. Thus, high frequency sound patterns appear to support the acquisition of both word form and meaning.

Frequency effects have also been observed in other languages, and Edwards and Beckman (2008) document these effects in two- and three-year-old children acquiring Cantonese, English, Greek, and Japanese. Recently, Másdóttir and Stokes (2016) observed robust correlations between frequency and production accuracy in two- and three-year-olds acquiring Icelandic. In sum, numerous studies have documented a dependency between expressive speech sound development on the one hand, and the frequency of sounds in the ambient language on the other hand. The dependency between frequency and production accuracy must be more narrowly defined, however, as the literature above could be true because children hear frequent sequences more often, because they produce them more often, or due to a combination of perceptual experience and production practice.

The relevance of production practice notwithstanding, the study of statistical learning has, to date, been primarily concerned with perceptual learning, and we will maintain our focus there. As such, when considering the literature on frequency effects in child speech development, the statistical learning framework suggests a basic hypothesis. Namely, if we control the frequency that children are exposed to a sound sequence *within an experiment*, we could see production accuracy change, depending on those experimental frequencies. That is, experimentally-implemented perceptual frequencies could simulate the relationship between speech accuracy and broader language statistics; the same relationship that has been documented by Edwards and Beckman (2008), Munson (2001), Zamuner et al. (2004), and others. So, within an experimental setting, child speech accuracy could be subject to statistical learning. Perhaps surprisingly, however, statistical learning has rarely been used to study expressive speech sound patterns.

Relative to the infant literature, the literature examining statistical learning in older children and adults is smaller. Statistical learning has been applied to pattern learning in typical adults (Eidsvåg, Austad, Plante, & Asbjørnsen, 2015; Finley, 2013; Onishi, Chambers, & Fisher, 2002; Richtsmeier, 2011, 2016). Statistical learning has also been studied in both typical and impaired populations, including for both receptive learning (Evans, Saffran, & Robe-Torres, 2009) and expressive learning (Plante, Bahl, Vance, & Gerken, 2011; Richtsmeier, Gerken, Goffman, & Hogan, 2009; Richtsmeier, Gerken, & Ohala, 2011). Focusing on expressive speech development, Richtsmeier et al. (2011) examined statistical learning in typically developing four-year-olds. Children were familiarized with novel words that were described as the names of make-believe animals. The words contained target word-medial consonant sequences, for example, the /fp/ in/ neɪfpən/. Across three experiments, children's speech accuracy improved significantly following exposure to a combination of multiple talkers producing multiple related words (e.g., /mæfpəm/, /baɪfpəm/, and /gɪfpək/ for /fp/), but not from exposure to multiple talkers producing a single word (e.g., just/ mæfpəm/), or from exposure to single talkers producing multiple words. The authors concluded that statistical learning—incorporating multiple types of variability—can facilitate accurate speech production.

In a study of statistical learning that incorporated talker variability, Plante et al. (2011) applied a similar paradigm to children with language impairments. Four- and five-year-old children with and without language impairments were familiarized with novel words like /mæfpəm/ and /boktəm/. Children heard those words produced either once by a single talker or 10 times by multiple talkers. When they heard many examples of the words produced by many different talkers, both typically developing children and children with language impairments increased their accuracy and shortened their production latencies. Plante and colleagues therefore provide compelling evidence that talker variability specifically, and statistical learning more generally, may play a role in addressing language learning deficits (cf. also Evans et al., 2009; Plante et al., 2014).

To summarize, statistical learning has been observed to facilitate language learning at multiple developmental stages, first allowing infants to learn ambient language structures and subsequently allowing children to increase production accuracy for within-experiment high frequency patterns. Statistical learning might facilitate speech production in a variety of ways, for example, by