Research papers

Lateralized processes constrain auditory reinforcement in human newborns

Anthony J. DeCasper *, Phyllis Prescott

Department of Psychology, UNC-Greensboro, Greensboro, NC 27402, United States

ARTICLE INFO

ABSTRACT

We investigated operant sucking response learning in human newborns. Auditory reinforcers always occurred monaurally to see whether their potency differed between ears. Experiment 1 – we controlled the reinforcers, either intrauterine heartbeat sounds or unfamiliar speech, while infants chose which ear received it. Experiment 2 – we controlled the reinforcers and the receiving ear. Unfamiliar speech reinforced learning only if infants could use their right ear and heartbeats reinforced learning only if infants could use their left ear. Experiment 3 – we controlled the ear while infants chose between their mothers' vs. a stranger's voice and between their mothers' vs. a foreign language. The more familiar speech reinforced learning only if infants could use their left ear. We proposed reinforcers' potencies differed between ears because the newborn's auditory system, just like adult's, optimizes their perceptual clarity by left-lateralized processing of their rapid temporal variations and right-lateralized processing of their longer-lasting spectral characteristics.

1. Introduction

Sounds can reinforce operant learning in human newborns immediately after birth. Some sounds are potent reinforcers, e.g., maternal speech and intrauterine heartbeats, others are not, e.g., male speech and instrumental music. In general, reinforcer potency increases as the acoustic similarity between the reinforcer and a sound the infant frequently experienced before birth increases. Thus, factors that affect reinforcer potency include the acoustic characteristics of sounds transmitted to the fetal ear, their frequency of occurrence, and the operating characteristics of the pre and perinatal auditory system, as well. Among the operating characteristics are systematic functional asymmetries, from cochlea to cortex. Whether they affect a reinforcer's potency is unknown. After a selective review of the relevant factors, we present three operant learning experiments that examined whether the functional asymmetries of the auditory system affect reinforcer potency.

1.1. Acoustic characteristics of the fetal environment

Sounds arising from maternal gastro-intestinal, cardiac, and vocal activity are ubiquitous (Lecanuet and Granier-Deferre, 1993). Acoustic waves generated by maternal activities propagate to fetal ear through tissue and bone without much distortion (Granier-Deferre et al., in press; Oliver, 1989; Querleu et al., 1989; Richards

et al., 1992). Near the placenta, intrauterine background noise is ~50 dB SPL at 300 Hz, falls off to near 30 dB SPL ~1 kHz, and reaches 0 dB SPL ~1800 Hz. Farther from the placenta, intensity is lower and the drop off over frequency steepens, e.g., reaching ~0 dB SPL ~1 kHz. Maternal heartbeat sounds exceed background noise by about 25 dB SPL. At 72 bpm the heartbeat sound consists of a sequence of approximately square wave pulses, 3 or 4 spectral bands between ~200 and 800 Hz that last ~200–300 ms, which recur at ~580-ms intervals. Intrauterine levels of maternal conversational speech exceed background noise by ~24 dB SPL. Her speech can be characterized as a non-random sequence of "complex sounds of varying temporal grain, including, periodic and aperiodic components, noise, frequency, and amplitude modulations" (Zatorre et al., 2002, p. 38). External speech recorded in utero is less intense because of attenuation by maternal tissue, with attenuation increasing along with frequency. Between 32-weeks gestational age (GA) and birth, the developing auditory system will be exposed to a stream of more than 5×10^6 maternal heartbeats and about 6.7×10^5 words spoken by the mother, most of them occurring in normal conversational utterances averaging about 10 words in length (Mehl et al., 2007; Rayson et al., 1997). Such is the acoustic environment of the late-term fetus.

1.2. The operating characteristics of the perinatal auditory system

1.2.1. Cochlea

Initial sound processing occurs at the cochlea, where activity of the outer hair cells produces audible otoacoustic emissions. A pair of tones lasting >~100 ms evokes distortion product otoacoustic

---

* Corresponding author. Tel.: +1 33 09 54 31 50 43.
 E-mail address: decasper@uncg.edu (A.J. DeCasper).

emissions (DPOAE) (Sininger and Cone-Wesson, 2004). DPOAE amplitude reflects the output of the cochlear amplifier. DPOAEs are adult-like around 38–40-weeks GA, especially with frequencies below 1500 Hz (Abdala, 1996; Aidan et al., 1997). Significantly, DPOAE amplitudes are greater at the left ear than at the right ear (Sininger and Cone-Wesson, 2004). A train of brief, rapidly presented clicks, e.g., 80 ms duration pulses presented at 89 Hz, is used to elicit transient evoked otoacoustic emissions (TEOAE). TEOAE reflects the output of the cochlear amplifier in response to brief duration, wide band stimuli. TEOAE amplitudes vary according to the signal's intensity, duration, rate of alternation, and spectra. Significantly, TEOAE amplitudes are greater at the right ear than at the left ear (Aidan et al., 1997; Sininger and Cone-Wesson, 2004).

1.2.2. Pathways

The spectral and temporal information transduced by the cochlea is transmitted via cochlear nerves to the cochlear nucleus. Throughout gestation the left cochlear nerve is morphologically more mature than the right, its axons are larger and more myelinated (Ray et al., 2005). At birth, it is functionally more efficient; latency of the auditory brainstem response (ABR) from the cochlea to the cochlear nucleus is shorter and its amplitude greater than for the right nerve (Eldredge and Salamy, 1996). Thereafter, ascending pathways partially cross. From the cochlear nucleus onward, ABR latencies are shorter and their amplitudes are greater on the right side than on the left (Eldredge and Salamy, 1996; Sininger and Cone-Wesson, 2006). Significantly, contralateral pathways are more numerous and their input to the cortex functionally dominates input from ipsilateral pathways (Witelson, 1987).

1.2.3. Cortex

The fetal cortex can process long lasting spectra and rapid variation. MEG recordings indicate the maternal heartbeat evokes cortical activity (Porcaro et al., 2006), and that two 100-ms tones, 500 Hz vs. 750 Hz, can be discriminated as early as 28-weeks GA (Draganova et al., 2007). Fetal cardiac activity also reflects stimulus discrimination. Cardiac measures indicate late-term fetuses can discriminate complex sounds based on their spectral differences such as male vs. female voices (Lecanuet et al., 1993), one low-pitch musical note vs. another note (Lecanuet et al., 2000), and the order of CV syllables, [ba][bi] vs. [bi][ba] (Lecanuet et al., 1987). They also discriminate their mother's voice vs. a stranger's, independently of what they are saying, (Kisilevsky et al., 2003), and a speech passage the mother had often recited vs. one she had never recited, independently of the voice of the speaker (DeCasper et al., 1994). Familiar passages can be discriminated as early as 32-weeks GA (Krueger et al., 2004).

Fetal responses to brief rapid variation has been shown to occur at 33-weeks GA, when a stream of alternating ~58 ms tones evoked an fMRI BOLD signal, but only in the left temporal area (in 2 of 3 fetuses) (Jardri et al., 2008). Recently, Granier-Deferre et al. (in press) stimulated late-term fetuses with a spoken sentence and with the same sentence in which white noise replaced all spectral information but retained the precise temporal variations. (The altered sentence sounds like rhythmic spattering of water on a hot griddle.) The normal and altered sentences evoked indistinguishable cardiac responses. The authors concluded that the processing of the rapid temporal variation in continuous speech can occur independently of the processing of its, longer lasting, spectral variation.

At birth, event related potentials (ERP) indicate discrimination between a standard 100-ms complex tone and a shorter one (Cheour et al., 2002), between maternal and non-maternal voices (deRegnier et al., 2000), and between CV vowels that last 250 ms, e.g., /gu/ vs. /ga/ (Alexandra et al., 2007). Significantly, ERPs from the right temporal area are most responsive to each of these differences in spectra. In contrast, Alexandra et al.'s (2007) study with newborns also found that ERPs that discriminated consonant-vowel transitions occurring within the first 40–45 ms of the CVs, e.g., /du/ vs. /gu/ were larger from the left temporal area. In a study somewhat analogous to that of Granier-Deferre et al. (in press), Peña et al. (2003) found, with optical topography, that left temporal areas of neonates showed significantly more activation to normal speech than to backward speech, i.e., normally ordered temporal variations in their natal speech vs. reversed variation.

In sum, the perinatal auditory system can process the rapid temporal variations and the spectral characteristics of a sound, including speech. Oftentimes the processes are found to be functionally asymmetric. The right cochlea/left auditory cortex differentially resolve rapid temporal variation, especially amplitude variation, occurring over brief intervals, <~40 ms, while left cochlea/right cortex differentially resolve spectral information occurring over longer intervals, >~100 ms. Whether the functional asymmetries have any effect on auditorially reinforced operant learning is unknown.

1.3. Auditorially reinforced learning in newborns

All operant learning studies conducted to date with newborns have presented the reinforcing sounds binaurally. The intrauterine heartbeat is a potent reinforcer (DeCasper and Sigafoos, 1983; Panneton and DeCasper, 1984). So are solo female singing (DeCasper, 1978), and mixed-voice singing with and without string accompaniment (Butterfield and Siperstein, 1972). The string accompaniment alone is not. The two other sounds shown to be ineffective are whispered female speech (Spence and Freeman, 1996) and male speech (DeCasper and Prescott, 1984). None of the non-reinforcing sounds is likely to occur in utero under normal conditions.

By far the most studied reinforcing sounds are various versions of continuous female speech. When used in choice studies, which allow the newborn to choose between two reinforcers, the effect of specific prenatal experiences become apparent.[1] The most potent reinforcer is the mother's voice that has been low-pass filtered at 500 Hz. It is more reinforcing than mother's voice low-pass filtered at 1 kHz or unfiltered (Moon and Fifer, 2000; Spence and Freeman, 1996). Mother's voice is more reinforcing than an unfamiliar voice (DeCasper and Fifer, 1980; Fifer and Moon, 1989). An unfiltered unfamiliar voice is more reinforcing than a low-pass filtered unfamiliar voice (Spence and DeCasper, 1987). Maternal language speech is more reinforcing than foreign language speech (Moon et al., 1993) and utterances/melodies the mother often recited/sang before birth are more reinforcing than novel utterances/melodies, independently of the voice itself (Cooper and Aslin, 1989; DeCasper and Spence, 1986).

The array of learning studies, all using binaural reinforcement, shows that prenatal auditory experience affects postnatal reinforcer potency. The effects of experience with spectral features of maternal speech, especially the lower frequencies, are clear. Three newborn learning experiments, presented below, assessed

---

[1] The specific transitive inferences of reinforcer potency, i.e., 1 > 2 > 3, permitted by the nine operant choice studies that were done are 1. [maternal voice, maternal language, repeated passage, unfiltered] ≈ [maternal voice, maternal language, novel passage, low-pass filtered at 500 Hz]; 2. [maternal voice, maternal language, novel passage, unfiltered] ≈ [maternal voice, maternal language, novel passage, low-pass filtered at 1 kHz]; 3. [non-maternal voice, maternal language, repeated passage, unfiltered]; 4. [non-maternal voice, maternal language, novel passage, unfiltered]; 5. [non-maternal voice, non-maternal language, novel passage, unfiltered]; and 6. [non-maternal voice, maternal language, novel, low-pass filtered at 1 kHz] ⩾ [non-maternal voice, maternal language, novel passage, low-pass filtered at 500 Hz] ⩾ [maternal voice, maternal language, novel passage, whispered speech] ≈ [non-maternal voice, maternal language, novel passage, whispered speech] ≈ [paternal voice, maternal language, novel passage, unfiltered] ≈ [non-paternal voice, maternal language, novel passage, unfiltered] ≈ zero.