INTRODUCTION

In the process of becoming a speaker of a language, one must acquire a tacit understanding of the phonological principles inherent in the native language. Phonological principles specify both phonetic variations that signal differences in meaning (phonemic distinctions) as well as systematic differences in production that occur within a language but do not signal differences in meaning. It is now well established that infants begin the process of acquiring the phonology of the native language during the first year of life (Jusczyk, 1997; Werker and Pegg, 1992). Of particular interest to our work, numerous studies show that in the early months of life, infants discriminate both native and non-native contrasts with equal ease, but 10- to 12-month-old infants perform like adults and only easily discriminate those contrasts that are used to differentiate meaning in their native language (Best et al., 1995, 1988; Kuhl et al., 1992; Polka and Werker, 1994; Trehub, 1976; Werker, 1989; Werker and Lalonde, 1988; Werker and Tees, 1984a).

The changes that occur in perception of non-native contrasts do not constitute an “absolute loss” of discriminative capacity. In tasks that mimic the processing demands of language comprehension, adults show an overwhelming bias to discriminate only that phonetic variation that carries phonemic status. However, when tested under other conditions, it is clear that the human auditory system retains the ability to discriminate non-native phonetic differences (for reviews see Strange, 1995; Werker, 1994). Thus the age-related decline in the ability to discriminate non-native contrasts is referred to as a functional (linguistic) reorganization rather than an absolute loss (Werker, 1995). The assumption has been that the functional reorganization involves a mapping of phonetic variation on to potentially meaningful differences.

The conclusion that the functional reorganization reflects sensitivity to potential phonemic contrasts rather than just sensitivity to language-specific phonetic variation is premature, however. The major problem is that much of the above work comparing discrimination of native and non-native contrasts has confounded linguistic exposure and phonological status (see MacKain and Stern, 1982, for a discussion). Thus we did not know whether infants, during the first year of life, maintain sensitivity to native contrasts simply because they have heard that phonetic variation occurring systematically in the language input, or whether they maintain sensitivity to only that phonetic variation that corresponds to meaningful distinctions in the native language.

In most previous work, a non-native contrast has typically been defined as a pair of syllables differing in only a single consonant (or vowel) that would signal a contrast in meaning in an unfamiliar language but not in the native language. In these experiments, however, researchers have failed to systematically control whether variants of one or both members of the non-native contrast also occur in the native language (but see Best et al., 1988; Polka, 1991). For example, in Werker’s research, one could argue that the initial position Hindi dental consonant does occur in English, but the initial position Hindi retroflex consonant does not (except in some speaker’s consonant clusters) in comparison
to the glottalized uvular and velar consonants from Nthlkanpx, neither of which occurs in English (Werker and Tees, 1984b). Without proper control of these factors, it is impossible to confirm whether or not the functional reorganization really does involve a mapping onto meaning. 

There are two ways to test whether the simple exposure hypothesis is sufficient to explain the developmental reorganization in phonetic perception. The first way to test this hypothesis is to ascertain whether complete lack of exposure necessarily leads to a decline in discrimination. The second way to test this hypothesis involves examining whether presence of exposure in the absence of phonemic status necessarily leads to maintenance of discrimination.

It has been shown that simple lack of exposure does not necessarily lead to a decline in discrimination performance on non-native contrasts. Best and colleagues (Best et al., 1988, 1995) tested English adults and English-learning infants of 8 to 14 months of age on a Zulu apical/lateral click contrast. This non-native contrast is comprised of phones which do not occur in English, and are not similar in any respects to the phonology of English. All subjects easily discriminated this contrast. Best and colleagues concluded that discrimination was maintained because the non-native phones were not assimilable to English phonology, and thus were spared the reorganization process.

Although Best et al.’s research (1988) reveals that lack of linguistic exposure may not always lead to a reduced ability to discriminate, we did not yet know if presence of linguistic exposure independent of phonological status is sufficient to maintain an ability to discriminate consonant contrasts. To conduct such a study, it is necessary to present listeners with a pair of phones to which they are routinely and regularly exposed (i.e., native) but that do not constitute a phonemic contrast in their native language.

English allows two alveolar stops in initial position—the voiced unaspirated stop, and the voiceless aspirated alveolar stop, [tʰ]. A third variant occurs following /s/—the voiceless unaspirated alveolar stop [t]. Linguistically speaking, [tʰ] and [t] are allophonic variants of the phoneme /t/ while [d] and [d] are allophonic variants of a single phoneme but rather come from two different underlying phonemic categories, /d/ and /t/. Thus, phonologically, [t³] and [tᵢ] are more similar to one another than either is to [d]. However, [tᵢ] and [d] are acoustically and perceptually more similar to one another than either is to [t³] (Lisker and Abramson, 1970; Lotz et al., 1960). Furthermore, although there are many languages in the world with a voiced unaspirated [d] and a voiceless unaspirated [tᵢ] phonemic contrast, it is not known whether there are any languages that contrast [d]s and [tᵢ]ṣ with the precise acoustic characteristics of the stimuli used in this study (see first Method section for description of the acoustic values).

To assess discrimination of a language specific phonetic difference using this set of phones, it was necessary to remove one of the context-dependent variants from its standard context and to present it in the context of the other. In previous work, Repp and Lin (1989) presented subjects with the [ba]–[pa] phonemic distinction but with both syllables preceded by an [s]. Adult English listeners were unable to perform the discrimination. This raises the possibility that the presence of [s] interferes with detection of phonetic information possibly due to masking. Thus, we decided to test English adults on the [tᵢ]–[d] phonetic difference and, to do this, we removed [tᵢ] from its standard [stᵢ] context (in English), and presented both [tᵢ] and [d] in syllable initial position.

The phones [tᵢ] and [d] are both components of English phonology and all English speakers are exposed to both [tᵢ] and [d]. However, although [tᵢ] and [d] are described formally as derived from two different underlying phonemes, the precise acoustic-phonetic difference between [tᵢ] and [d] in syllable initial position is never used to contrast meaning in English. Thus, by assessing discrimination of [tᵢ] vs [d], we can address the question of whether linguistic exposure lacking in phonemic status is sufficient to allow maintenance of discrimination.

To address this main question, adults and both 6- to 8- and 10- to 12-month-old infants were tested in the conditioned head turn procedure. It is in this category change procedure that much of the evidence of a developmental reorganization in phonetic perception has previously been revealed. If exposure is the main factor in the developmental reorganization, then listeners in all age groups should easily discriminate [tᵢ] from [da] because they are regularly and routinely exposed to this phonetic variation, albeit not both in initial position. If, on the other hand, phonological status plays a major role, then both adults and 10- to 12-month-old infants should have difficulty discriminating this nonmeaningful but native phonetic variation.

Before addressing the main question, it was necessary to examine English adults’ perception of these stimuli in detail. In experiment 1, adults were tested in a procedure designed to determine if both [tᵢ] and [da] are perceived as members of the same English phoneme or as members of two different phonemes. To do this, adults were tested in a variation of a category goodness task (Miller, 1995). Next, to determine if these segments are discriminable under sensitive testing conditions, adults in experiment 2 were tested in a same/ different (AX) task with a 500-ms ISI. The AX procedure with a short ISI is thought to be sensitive because adults tested in this procedure discriminate some difficult non-native contrasts (Werker and Logan, 1985) and also discriminate difficult within consonant phonetic variation (Carney et al., 1977). Thus adults tested in an AX procedure may discriminate these consonants even if both segments are judged to be members of the same phonemic category.

I. EXPERIMENT 1

A. Introduction

The purpose of the first experiment was to determine if English adults judge the English syllable [tᵢ] as linguistically equivalent to the syllable [da], or, alternatively, detect subtle differences between them. Because of the overwhelming acoustic similarity of [tᵢ] and [da], it was predicted that English listeners would identify all exemplars as members of an English /d/. It was possible, however, that adults would identify differences in quality between [tᵢ] and [da]. Con-
sistent acoustic differences between these phones do exist, albeit not cues to a phonemic contrast (see Table I). This detail could allow at least some listeners to detect differences in quality between [t\textsuperscript{a}] and [da]. Specifically, because [t\textsuperscript{a}] never occurs in initial position and yet it is being presented in initial position, adults could have judged [da] to be a better example of the category of initial position alveolar stops than [t\textsuperscript{a}].

One recent study investigated adult English speakers’ discrimination of two English allophones presented in their natural context and an unnatural context (Whalen et al., in press). In English, syllable initial “L” differs from syllable final “L.” Initial position “L” is transcribed phonetically as light [l], and final position “L” is dark [l]. Whalen et al. hypothesized that adults would judge an allophone presented in the appropriate position as a better example of the category “L” than the allophone presented in the inappropriate position. Accordingly, they presented adults with dark and light “L”’s in both initial position and final position, [l, l, el, el], and care was taken to equate the vowels. When tested in a category goodness task, adults did not judge allophones occurring in their appropriate position, [l] and [el], as better examples of the category than those allophones occurring in the inappropriate position, [el] and [el]. All exemplars were considered good examples of the category. In addition, when tested in an AXB procedure, adults did not exceed chance levels of discrimination. Thus in their study, English adults were not able to detect differences between the English allophonic varieties of dark and light “L” in either an identification or a discrimination task.

The stimuli used in our studies differ from the Whalen et al. (in press) stimuli. Whereas light [l] and dark [l] are allophones of the same underlying consonant, the syllables [t\textsuperscript{a}] and [da] arise from different underlying consonants. Thus, even though adults do not detect differences between [l] and [l] it is possible that adults will detect differences in category goodness between [t\textsuperscript{a}] and [da].

A variation of a procedure designed by Miller (1995) was used to try to elicit differential ratings of quality among the phonetic exemplars. Adults were given a sheet with one syllable printed at the top (either DA, STA, or SDA). SDA was included because it was thought that subjects might attend to phonetic differences among the stimuli when comparing them to a phonologically and orthographically impossible syllable. They were asked to rate the quality of match of individually presented exemplars of [t\textsuperscript{a}] and [da] to the letter printed in bold. If all exemplars are assimilated to the meaningful /d/ category, ratings from subjects with DA page headings should be better than ratings for those with STA and those with SDA should fall in the middle. Differences in quality will be evident if adults rate [da] exemplars as better examples of DA than [t\textsuperscript{a}] and conversely, rate [t\textsuperscript{a}] as better examples of STA than [da].

### B. Method

#### 1. Subjects

Thirty university students between the ages of 18 and 30 years participated in this study. All subjects were monolingual speakers of English and had no knowledge of any other language before the age of 8 years.1 Subjects initially indicated interest in participating in research by adding their name to a subject pool. They were then contacted by phone, given details of the procedure, and if they were interested, an appointment was made. Following the procedure, they were given a verbal description of the study and a two-page summary explaining the theoretical basis of the research project. Subjects were also given 1 credit for an undergraduate psychology course.

#### 2. Stimuli

The speech samples used in this study were produced by a male, native English speaker. Several exemplars of [da] and [st\textsuperscript{a}] were digitized directly into a Macintosh II FX computer using the Signalize speech analysis program via a GW Instruments analog-to-digital board (model GWI-AMP). In addition, several exemplars of [t\textsuperscript{a}a] were recorded. The [t\textsuperscript{a}a] exemplars were not used in experiments 1 and 2, but were used in experiment 3. All signals with perceived differences in cues such as pitch, loudness, intonation contour, and duration were eliminated.2 Fifteen exemplars were selected—five for each category. Finally, the selected [st\textsuperscript{a}] exemplars were modified by deleting the initial [s], taking care not to remove any of the characteristics of the burst release of the alveolar stop consonant.

The five exemplars selected for each [da], [t\textsuperscript{a}], and [t\textsuperscript{a}a] category were analyzed to identify criterial phonetic characteristics. As can be seen in Table I, the phonemes [da] and [t\textsuperscript{a}a] differ in VOT, burst duration, fundamental frequency (F0) at vowel onset, F1 at vowel onset (the first formant), and F2 at vowel onset. In contrast, the syllables [da] and [t\textsuperscript{a}] differ only in F2 onset (possibly due to coarticulation of the preceding /s/) and have overlapping values in each of the other measures (see also Repp and Lin, 1987). The main acoustic cue for differentiating [t\textsuperscript{a}] from [da] is, therefore, a higher F2 onset for [da] than [t\textsuperscript{a}]. It should be noted that it has been reported in the literature that the mean

### Table I. The results of acoustical analyses of the exemplars.

<table>
<thead>
<tr>
<th>Item</th>
<th>VOT (ms)</th>
<th>Burst duration (ms)</th>
<th>Syllable duration (ms)</th>
<th>F0 onset (Hz)</th>
<th>F1 onset (Hz)</th>
<th>F2 onset (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t\textsuperscript{a}1</td>
<td>17.45</td>
<td>23.60</td>
<td>336</td>
<td>106</td>
<td>675</td>
<td>1479</td>
</tr>
<tr>
<td>t\textsuperscript{a}2</td>
<td>41.00</td>
<td>22.90</td>
<td>397</td>
<td>105</td>
<td>726</td>
<td>1366</td>
</tr>
<tr>
<td>t\textsuperscript{a}3</td>
<td>35.75</td>
<td>20.45</td>
<td>389</td>
<td>103</td>
<td>726</td>
<td>1400</td>
</tr>
<tr>
<td>t\textsuperscript{a}4</td>
<td>46.50</td>
<td>34.50</td>
<td>407</td>
<td>102</td>
<td>718</td>
<td>1298</td>
</tr>
<tr>
<td>t\textsuperscript{a}5</td>
<td>36.10</td>
<td>33.75</td>
<td>415</td>
<td>105</td>
<td>675</td>
<td>1358</td>
</tr>
</tbody>
</table>

*da*1 | 6.30 | 15.20 | 350 | 96 | 449 | 1600 |
| da2 | 5.30 | 12.95 | 374 | 101 | 476 | 1600 |
| da3 | 3.50 | 14.60 | 353 | 98 | 458 | 1643 |
| da4 | 3.35 | 14.05 | 403 | 98 | 432 | 1617 |
| da5 | 6.20 | 13.05 | 369 | 100 | 458 | 1505 |

*ta*1 | 7.55 | 14.85 | 347 | 101 | 441 | 1410 |
| ta2 | 2.95 | 15.60 | 338 | 100 | 458 | 1418 |
| ta3 | 1.70 | 15.25 | 353 | 100 | 459 | 1410 |
| ta4 | 6.35 | 17.45 | 339 | 100 | 484 | 1436 |
| ta5 | 3.75 | 15.55 | 362 | 101 | 484 | 1401 |
F0 onset of \( [\mathrm{t}^-\mathrm{a}] \) and \([\mathrm{da}]\) differ. Indeed, lower mean F0 onset for \([\mathrm{da}]\) is reported to be one of the consistent acoustic cues differentiating these allophonic variants (e.g., Hombert et al., 1979). Although there is a mean difference in values of F0 at vowel onset for our stimuli \([\langle [\mathrm{da}]\rangle = 98.6 \text{ vs } [\mathrm{t}^-\mathrm{a}] = 100.4\)\), F0 onset is not a distinguishing cue since there is overlap in F0 onset between the two categories.

The ten digitized stimuli in the stimulus set (five \([\mathrm{t}^-\mathrm{a}]\)s and five \([\mathrm{da}]\)s) were imported into the Bliss Experimental Control System. This program controlled the presentation of the stimuli.

3. Apparatus and procedure

Subjects were given a response sheet with either DA, STA or SDA printed at the top (ten subjects in each condition). Subjects were presented with the ten individual exemplars and their task was to rate the goodness of each of the exemplars in relation to the referent, the one printed in bold at the top of their page. Each individual exemplar was followed by a 5-s response interval during which subjects wrote a number between 1 (poor) and 7 (excellent) on the response sheet.

Subjects were seated in a sound-attenuated chamber in groups of no more than three. They were seated with their backs to each other to ensure that no subject saw the other’s response sheet. Before being given the response sheet, subjects were told not to utter the syllable on the top of the page and were given instructions to rate the quality of each presented signal relative to the referent shown in bold print at the top of their page. They were also told that although the task was very difficult for most people, they should nevertheless try to listen carefully for differences between the sounds. It was hoped that such explicit instructions would facilitate detection of subtle quality differences. Exemplars were presented free field by a Compac 286 computer using the Bliss Program and were presented at a comfortable listening level of 65 dB over a Bose speaker.

Subjects were tested in two sessions with a 3–5-min break between sessions. Each of the ten individual exemplars was presented six times per session. Thus, in total there were 60 presentations of exemplars from the category \([\mathrm{da}]\) and 60 from the category \([\mathrm{t}^-\mathrm{a}]\) for a total of 120 responses, 12 judgments for each individual exemplar. The order of presentation was randomized for each group of subjects and for each session.

C. Results and discussion

To determine if judgments of category goodness differed as a function of the category of the phone, a mixed-model analysis of variance (ANOVA) was conducted in which the group variable was page heading (STA, SDA, or DA) and the two repeated measures were session (session 1 versus session 2) and phone category ([\mathrm{da}] vs [\mathrm{t}^-\mathrm{a}]). The dependent variable was the rating value.

There was a main effect for page heading \([F(2,27) = 5.47, p < 0.01]\) and an interaction between session and page heading \([F(2,27) = 3.24, p < 0.05]\). No other main effects or interactions were significant. The main effect for page heading was analyzed by comparing mean rating values using Fisher’s Protected LSD. The results indicated that adults judged all exemplars as less good when STA was the page heading than when either DA \((p < 0.005)\) or SDA \((p < 0.02)\) were the headings (see Fig. 1). Ratings were not significantly different when SDA versus DA were page headings.

The interaction between session and page heading was analyzed by comparing session for each page heading. When STA was the page heading, session was not significant. Likewise, when DA was the page heading, session was not significant although the means were slightly higher during the second session \((M = 4.71, \text{ s.d.} = 0.92; \text{ second session } M = 4.85, \text{ s.d.} = 1.00)\). In contrast, when SDA was the page heading, there was a trend for higher goodness ratings during the first session \([F(1,9) = 3.87, p < 0.08]\) \((\text{first session } M = 4.74, \text{ s.d.} = 0.42; \text{ second session } M = 4.35, \text{ s.d.} = 0.77)\).

Still, during the second session, signals were not necessarily considered poor examples of the category, simply less good.

This study also provided important information as to whether any of the ten individual exemplars in the stimulus set were perceived as consistently different in quality. A mixed model ANOVA was conducted using page heading as a group variable and exemplar as a repeated measure. This analysis of mean rating scores revealed that although page heading was significant \([F(2,27) = 5.80, p < 0.008]\), and there was an interaction between page heading and exemplar \([F(18,243) = 2.75, p < 0.02]\) \((\text{Greenhouse-Geisser adjustment})\), there was no main effect for exemplar. This suggests that all exemplars fell within an acceptable range and, therefore, all exemplars were used in subsequent experiments.

Taken together, these results support the prediction that adults judge all exemplars as better examples of ‘‘D’’ than ‘‘T’’. Interestingly, even though English spelling rules do not allow ‘‘SD,’’ listeners also rated the signals as good examples of the SDA referent. They did not, however, rate the exemplars as good examples of the STA referent even though English spelling rules allow ‘‘ST.’’ Indeed, inspection of the means for ratings of the individual exemplars shows that even the highest rated exemplar for ‘‘ST’’ was rated as less good than the lowest rated exemplar for either DA or SDA. Thus, the most important fact arising from these results was that both \([\mathrm{t}^-\mathrm{a}]\) and \([\mathrm{da}]\) were assimilated to the English phoneme ‘‘D’’.

There was no indication that adults judged the phone in its appropriate context, \([\mathrm{d}]\), as a better example of the category than the phone in its inappropriate context, \([\mathrm{t}^-\mathrm{a}]\). This
replicates the results from Whalen et al. (in press) suggesting that native speakers do not detect allophonic variation. The present study goes one step further, however, in that even when the phonetic variation arises from two different underlying meaningful categories specified by English phonology, adults do not detect differences among phones.

II. EXPERIMENT 2

A. Introduction

The second step in assessing adult perception of a language-specific phonetic difference was to determine if adults can discriminate \([\text{da}]\) from \([\text{t}^\text{a}]\) under relatively sensitive testing conditions. Even though adults did not detect quality differences in the signals in the first study, it was still possible that adults would discriminate \([\text{da}]\) from \([\text{t}^\text{a}]\) if tested in a more sensitive procedure. The AX (same–different) procedure has been shown to facilitate discrimination of phonetic variation that is not meaningful in the native language (Carney et al., 1977). In this procedure, adults are presented with a pair of speech sounds and are asked to say whether the two speech sounds are the same or different. As mentioned previously, ease of discrimination can differ as a function of the ISI between the pair of stimuli (Werker and Logan, 1985). Thus, to facilitate discrimination of the phonetic difference, adults were tested in the AX procedure using a 500-ms ISI.

Order was controlled in this study to allow for the possibility that adults would show better discrimination in one order than in another. Previous research in vowel perception has shown that order can affect discriminability. In particular, when infants are tested in the conditioned head turn procedure (Kuhl et al., 1992; Polka and Werker, 1994) and adults are tested in an AXB procedure (Sussman and Lauckner-Morano, 1995) discrimination is worse in one order than the other. Kuhl et al. (1992, 1993) proposes a perceptual magnet effect to explain this phenomenon. According to the magnet effect, discrimination will be worse if the more prototypical vowel is presented as the referent than if the nonprototype is the referent because the more prototypical vowel acts as a “magnet,” pulling the others toward it. In Kuhl’s work, subjects were always presented with good and poor instances of a single phonetic category (i.e., good and poor /i/ and /a/). In an extension of this work, Polka and Werker (1994) presented subjects with two different German vowels, both of which were perceived by English listeners as instances of a single English vowel category (see also Polka and Bohn, 1996). Similarly, as revealed in experiment 1, both \([\text{da}]\) and \([\text{t}^\text{a}]\) are perceived as instances of the single English phoneme category /da/ even though they are different phonetically and arise from two different underlying phonemes. Thus, to allow for an analysis of a possible magnet effect with our stimuli, order was controlled.

B. Method

1. Subjects

Eighteen adults between 18 and 25 years of age participated in this study. They were recruited as described in experiment 1 and the selection criteria were identical. Two adults’ data were excluded (one due to a self-reported hearing loss in one ear and one due to equipment error), resulting in a sample size of 16.

2. Stimuli

The ten stimuli used in this study were the five \([\text{t}^\text{a}]\)s and the five \([\text{da}]\)s described in experiment 1. The set of pairings included each individual exemplar paired with every other exemplar. This resulted in five different (DIFF) pairs and four same (SAME) pairs for each exemplar (e.g., DIFF \(=\{\text{da}\}1 \text{ with } \{\text{t}^\text{a}\}1, \{\text{t}^\text{a}\}2, \{\text{t}^\text{a}\}3, \{\text{t}^\text{a}\}4, \text{ and } \{\text{t}^\text{a}\}5\) ). SAME \(=\{\text{da}\}1 \text{ with } \{\text{da}\}2, \{\text{da}\}3, \{\text{da}\}4, \text{ and } \{\text{da}\}5\). Because physically identical pairings may alter the response bias of a subject (Werker, 1993; Whalen et al., in press) no exemplar was paired with itself. Thus, one additional SAME contrast was randomly selected for each exemplar to equalize the numbers of SAME and DIFF pairs.

3. Apparatus and procedure

Presentation of contrasts was controlled by a 286 Compaq computer using the Bliss Program. Pairs were presented free field in an IAC sound attenuated chamber over a Bose speaker with a 500-ms ISI and a 2000-ms response interval. During the response interval, subjects pressed buttons marked SAME or DIFF on a mouse connected to the computer. The computer recorded all responses.

Adults were tested individually in three stages: a familiarization stage, and two testing stages. To familiarize adults with the procedure, they were presented with 16 pairings alternating between two SAME trials and then two DIFF trials (in the same order for all subjects). Each adult was given a list of the correct responses and was told to respond on each trial by pressing the appropriate button. After this brief familiarization, testing began.

Two rounds of testing were conducted with 100 trials in each set. Within each set, 50 pairs began with a \([\text{t}^\text{a}]\) and 50 began with a \([\text{da}]\). Of those 50, 25 were SAME pairs and 25 were DIFF. Each exemplar occurred ten times in the “A” position. To ensure that no adult heard the pairs in the same order twice or in the same order as another adult, the order of presentation of the 100 pairings was randomized for each round and for each adult.

A’ scores were calculated for each adult. A’ scores control for the response bias of a subject by taking into consideration the frequency of false alarms relative to the number of hits. This is a nonparametric statistic similar to d’ but, unlike d’, can be used with a smaller number of trials. In addition A’ values are constrained to vary between 1 (perfect performance) and 0.5 (chance performance). A value below 0.5 is mathematically possible but suggests that the subject is systematically using some cue or strategy different than that required for the task.

C. Results

To determine whether English listeners could discriminate the syllables, a repeated measures ANOVA was conducted. The repeated measures were session of testing (first versus second) and order of pairings ([da] first versus [t a])
There was a main effect for order \( F(1,15) = 36.87, p < 0.001 \) and an interaction between order and session \( F(1,15) = 7.24, p < 0.02 \). The main effect for order indicated that listeners had higher \( A' \) scores when \([t^a]a\) was in first position \((M = 0.75, \text{s.d.} = 0.12)\) than when \([da]\) was in first position \((M = 0.56, \text{s.d.} = 0.15)\).

Follow-up analyses of the interaction between order and session revealed a significant difference between the first and the second session only when \([da]\) was in first position \( F(1,15) = 8.68, p < 0.01 \) (see Fig. 2), but the magnitude of the order effect was still apparent even when taking into account the interaction. That is, although discrimination for \([da]\) first pairings improved from the first to the second session, discrimination for \([da]\) first pairings during the second session did not reach the levels of accuracy seen when \([t^a]a\) was in first position.

\( A' \) scores in each of the conditions were compared to chance and then to levels of discrimination expected for a native contrast (\( A' \) native discrimination is 0.95; see Polka, 1991). Importantly, when \([t^a]a\) was in first position, the \( A' \) scores were well above chance \((t(15) = 10.44, p < 0.0001; \text{session 2}, t(15) = 7.01, p < 0.0001)\), and yet below native levels \((t(1,15) = 7.39, p < 0.0001; \text{session 2}, t(15) = 6.02, p < 0.0001)\). The results differ dramatically when \([da]\) was in initial position. In this condition, the \( A' \) scores were greater than chance only during the second session \((t(15) = 3.078, p < 0.008, M = 0.60)\). As well, mean \( A' \) were still well below native levels of discrimination in both sessions \((t(1,15) = 11.73, p < 0.0001; \text{session 2}, t(15) = 10.35, p < 0.0001)\). Thus it appears that when \([da]\) was in first position, discrimination was attenuated.

The direction of the order effect is consistent with predictions from the perceptual magnet hypothesis (Kuhl et al., 1992). Discrimination is attenuated when the (putatively) better exemplar is presented as the referent. Further investigation revealed, however, that the magnet effect cannot explain our results. In her original work, Kuhl (1991, 1993) suggested that when a ‘‘good’’ exemplar from within a category is the referent, discrimination is attenuated because perceptual distance is reduced between the good example and less good examples. Thus by extension, we can deduce that stimuli compared to the good exemplar should be judged the SAME more often than stimuli compared to the less good exemplar. In the present case, since \([da]\) first pairings showed reduced discriminability (appeared to act as the perceptual magnet), adults should be responding SAME more often to pairs with \([da]\) first than with \([t^a]\) first. However, they were not. In a repeated measures ANOVA comparing percent of SAME responses as a function order of pairings \(([da] \text{ first versus } [t^a] \text{ first}),\) subjects were significantly more likely to respond SAME to \([t^a]a\) first pairings \((M = 53.19, \text{s.d.} = 22.61)\) than to \([da]\) first pairings \((M = 42.81, \text{s.d.} = 15.54),\) an effect opposite to predictions from the perceptual magnet effect.

D. Discussion

The results from this experiment show that English adults can discriminate \([t^a]a\) from \([da]\) when tested in a sensitive procedure with a 500-ms ISI. When \([t^a]a\) was in first position, adults correctly discriminated well above chance levels. That this was not an easy or straightforward task was evidenced by the fact that when \([da]\) was in first position, performance was severely attenuated particularly during the first block of 100 trials and by the fact that performance in all cases was still worse than would be expected on a native language phonemic contrast. Accuracy improved during the second session for those pairings with \([da]\) first and reached greater than chance levels of accuracy. Thus, discrimination was better than chance in three of four conditions. These results provide evidence that adults tested in a sensitive procedure can discriminate (nonmeaningful) native language phones, and can do so even when one of the phones is presented in a nonstandard context but that their discrimination levels are lower than would be expect for a native phonemic contrast.

These results differ somewhat from the results in Whalen et al. (in press). Two main differences between this study and theirs may help us to understand this difference. First, the phonetic difference in the present study is drawn from two different meaningful categories while the Whalen et al. stimuli are drawn from a single underlying phonemic category. As such, the present stimuli might be more easily discriminated, although underlying phonological status does not necessarily predict phonetic dissimilarity. Second, a more adequate explanation may be that the AX procedure used herein may be more sensitive to within category distinctions than the AXB procedure used by Whalen et al. (see Carney et al., 1977, for evidence). This difference in sensitivity between the two procedures may account for the different results. There was strong evidence of an order effect. When \([da]\) was in initial position, performance was attenuated and even though there was improvement during the second session, accuracy did not reach the levels of accuracy seen when \([t^a]a\) was in first position. The direction of the order effect was, at first blush, consistent with predictions from the perceptual magnet hypothesis (Kuhl et al., 1992). In contrast to predictions from the perceptual magnet hypothesis, however, adults were less likely to respond SAME to \([da]\) first pairings. The lack of support for a magnet effect was perhaps not surprising given the lack of differences in judgments of category goodness for \([t^a]a\) vs \([da]\) in experiment 1. According to Kuhl’s work (Kuhl, 1991, 1993) a magnet effect would only be predicted if there were differences in goodness be-
between the two sets of stimuli. Thus if the order effect found here is interpretable, some mechanism other than the magnet effect must be responsible. It may be that a more differentiated explanation that includes both contrast and anchor effects is required (see Warren, 1985).

III. EXPERIMENT 3

A. Introduction

The final and critical experiment was designed to assess discrimination of the English syllables [tʰa] and [da] by adults and by infants of both 6 to 8 and 10 to 12 months of age. The results from experiment 2 reveal that under sensitive testing conditions, English adults can discriminate [tʰa] from [da]. However, it was unclear if adults would discriminate this distinction when tested in a category change procedure. Previous research has shown that adults tested in a category change procedure easily discriminate only those consonant contrasts that are meaningful in the native language (Werker et al., 1981; Werker and Tees, 1984b). Thus it was not known whether the conditioned head turn procedure, as an instance of a category change procedure, would be sufficiently sensitive to reveal discrimination of nonmeaningful, native language phonetic variation.

As mentioned in the Introduction, most previous research has confounded exposure and phonological status. The stimuli used in the present research were selected specifically to unconfound these two factors. If the phonological status of a contrast is important in adult speech perception and the category change procedure elicits discrimination of only meaningful contrasts, adults will not discriminate [tʰa] from [da] in this procedure. On the other hand, if simple exposure is sufficient to maintain (or reestablish) discriminability, adults should discriminate these syllables even when tested in the category change task.

There were also two possible outcomes for the younger infants. On the one hand, previous evidence suggests that 6- to 8-month-old infants can discriminate almost every consonant contrast with which they have been tested (Best et al., 1995, 1988; Kuhl et al., 1992; Polka and Werker, 1994; Trehub, 1976; Werker, 1989; Werker and Lalonde, 1988; Werker and Tees, 1984a). On the other hand, the broad-based sensitivities shown by 6- to 8-month-old infants could reflect a sensitivity to only those phonetic differences that are used as meaningful contrasts in the world’s languages. As mentioned in the General Introduction, there are many languages that use voiced, unaspirated [d] versus voiceless unaspirated [tʰ] as a phonemic contrast but our particular stimuli include no systematic differences in VOT. Thus the particular combination of phonetic cues used in our English [tʰ] and [d] stimuli may not correspond to the phonetic cues in a meaningful contrast in any of the world’s languages.

Finally, there were two possible outcomes for the older infants. As mentioned previously, absence of exposure to a speech contrast does not always lead to decreased ability in 10- to 12-month-old infants because English-learning infants this age and older discriminate the non-English Zulu apical/lateral click contrast (Best et al., 1988). What we did not yet know was whether 10- to 12-month-old infants would discriminate speech segments to which they are exposed but which do not signal meaningful differences in the native language. The variants [da] and [tʰa] allowed us to test this possibility since these syllables are part of the native language and yet do not contrast meaning.

B. Method

1. Subjects

Twelve English monolingual adults between 18 and 22 years of age participated in the study. Recruitment was identical to experiment 1 and selection criteria were identical. Two additional subjects were also tested in this procedure but their data are reported separately because they had both received extensive phonetics training prior to participating in the study.

English-learning infants were recruited in several ways. Most parents whose babies participated were initially contacted by a research assistant from our laboratory who visited post-partum mothers in the hospital. The research assistant described the type of studies done in our laboratory in general terms and asked mothers if they were interested in participating in our studies. If the mother indicated interest, she provided her address and phone number and gave permission to be contacted at a later date. When the baby was the appropriate age, parents were contacted by phone and given detailed information about this study. If they were still interested, an appointment was made. Other infants were enrolled when their parents contacted the laboratory after seeing our posters or after hearing advertisements we had broadcast on local radio stations. Finally, some mothers and infants were recruited by parents who had participated in our research. These mothers called us to make appointments. All infants were healthy, born within 2 weeks of due date, were from English-speaking homes, and had not been exposed to any other language more than an estimated 10%. They were given a T-shirt and a certificate after participating.

Data from 32 infants were used: 20 English-learning infants aged 6 to 8 months (6 females, 14 males, mean age = 6 months and 28 days) and 12 English-learning infants aged 10 to 12 months (6 females, 6 males, mean age = 10 months 15 days). An additional 30 infants aged 6 to 8 months were excluded from the study due to failure to condition to the native meaningful contrast on the first visit (15), failure to recondition to the phonemic contrast on the second visit (4), crying or ill (6), equipment error (2), and failure to return for the second visit (3). Seventeen additional older infants were excluded due to failure to condition to the native meaningful contrast on the first visit (9), failure to recondition to the meaningful contrast on the second visit (2), cried or ill (5), and equipment error (1).

2. Stimuli

The stimuli were the ten exemplars described previously, the five [tʰa]s and five [da]s. The stimuli were presented on-line using a Data Translation 2801 A board. Syllables were presented free field over a Bose speaker at 65 dB with a 1500-ms ISI. The category change procedure was controlled by a 286 Compaq computer using custom software.
Since it was necessary to ensure that infants can and will perform in this task, infants were initially tested using an English phonemic contrast. We selected the phonemic distinction [tʰa] vs [da] as a control so that all discriminations would involve alveolar stops. This would ensure at least some measure of comparability between the control and experimental sessions. Thus, in addition to the five [tʰa] and the five [da] exemplars used in the adult studies, the five exemplars of the phoneme [tʰa] described in experiment 1 were included in the set of stimuli.

3. Apparatus and procedure

The testing was conducted in an IAC sound-attenuated chamber. Subjects were presented with a repeating background signal with a 1500-ms ISI. At random intervals, the signal changed for at least three presentations and then returned to the background signal. Thus subjects heard either [da] as background and [tʰa] as change or the reverse. The procedure for infants involves conditioning infants to turn their head when they detect a sound change (for more details see Polka et al., 1995; Werker et al., in press).

Infants were tested in three stages: familiarization, conditioning, and testing. During the first two stages, infants were presented with only one exemplar from each of the two categories of speech sounds. During the testing stage, infants were presented with several exemplars and thus must attend to categorical differences among the speech signals. During the first stage, every trial was a change trial. When the infant was judged to be in a state of readiness, syllables from the other category were presented and the reinforcer was activated immediately. This stage comprised five trials. During the second stage, the delay between the change in stimuli and activation of the reinforcer was lengthened, giving the infant the opportunity to make an anticipatory head turn and learn the contingency between a change in the speech signal and the activation of the visual reinforcers. Once an infant correctly performed three consecutive head-turns or the number of trials reached a maximum of 15, the third stage began. The third stage was the testing stage. The computer randomly selected control (no-change) or experimental (change) trials. During this stage, the randomization program was set such that change trials occurred on approximately 60% of the trials with the further restriction that no more than three consecutive control or three consecutive change trials could occur. The testing stage included 25 trials unless an infant was within two trials of reaching criterion (seven out of eight consecutive correct responses) at which time an additional five trials were presented. If an infant failed to respond on three successive change trials, retraining occurred in which the infant was presented with only single exemplars. During these retraining trials, if no head-turn occurred the reinforcer was automatically activated on the third change stimulus and two more presentations of the change stimulus were presented. After the three retraining trials, the program automatically returned to testing. Each infant was limited to a maximum of two retraining sets and these trials were not included in the analyses.

Infants were tested on two days: the first day they were presented with the English contrast [tʰa] vs [da] to ensure that they would perform in the task. Infants who failed to show evidence of discriminating [tʰa] from [da] were not tested on a second day. Those infants who successfully discriminated the syllables [tʰa] and [da] on the first day were presented with the syllables [tʰa] and [da] on the second day. If they failed to discriminate between these syllables they were retested with the [tʰa] vs [da] phonemic contrast to ensure that their failure was not due to forgetting the task. If infants failed to recondition on the phonemic contrast, their data were not included in the analyses.

The procedure for adults is very similar to that used for infants except adults are asked to raise their hand rather than turn their head when they detect a change from the referent to a different category (see Polka, 1995 for details). In our procedure, adults were only tested on the [tʰa] and [da] phones, unlike the infants who were also tested on the [tʰa] vs [da] phonemic contrast. As such, there were only two phases to the adult testing procedure, training and testing. During the training phase, subjects were presented with ten change trials to familiarize them with the procedure. If subjects raised their hand during a change trial, the reinforcer (a light) was activated. If a subject failed to respond to a change trial (a miss), the reinforcer was turned on automatically after the third presentation of the exemplar from the change category and two more presentations of the different exemplar occurred while the light remained on. Then the background exemplar was presented again. At the end of ten trials, the subjects were told that testing was about to begin. The test phase was otherwise identical for infants and adults.

Indicators of performance included number of subjects reaching criterion and A’ scores (see experiment 2).

C. Results

To allow comparison of 6- to 8- and 10- to 12-month-old infants, an analysis of proportions (ANPRO, Marascuilo, 1966) was conducted comparing the number of subjects in each group reaching the criterion of seven out of eight consecutive correct responses. The ANPRO is an analog of chi-square that allows comparison of unequal n’s and does not require a minimum number in each cell. The analysis revealed that a significantly greater proportion of 6- to 8-month-old infants reached criterion than did the proportion of 10- to 12-month-old infants ($\chi^2 = 11.62, p < 0.003$). There was no effect of order of presentation. Whereas 11 of the 20 younger infants reached criterion (6 when [da] was the referent and 5 when [tʰa] was the referent), only 1 of the 12 older infants did so (and this was when [da] was the referent).

To compare the infant to the adult data, a second ANPRO was conducted including adults, as well as 6- to 8- and 10- to 12-month-old infants. This analysis revealed a significant difference between the proportion of adults and the proportion of 10- to 12-month-old infants reaching criterion ($\chi^2 = 13.67, p < 0.001$), indicating that discrimination of the syllables was attenuated in 10- to 12-month-old infants (see Fig. 3). The difference between the proportion of adults and the proportion of 6- to 8-month-old infants reaching criterion was not significant.
can discriminate support for the assertion that both younger infants and adults parallels the results from the ANPRO and provides further distinction unlike their typical chance level performance in phonetic variation in the language. If this hypothesis is correct, all age groups should have discriminated the difference between [da] and [t’a] because this phonetic difference occurs in the linguistic input. According to the second hypothesis, the reorganization occurs as a function of the phonemic status of the phones in question. If this hypothesis is correct, only the younger infants, who as yet have no knowledge of the system of phonological contrasts in their native language, should have discriminated [t’a] and [da], and both older infants and adults should have failed to discriminate this pairing.

D. Discussion

The results from this study reveal that English infants aged 6 to 8 months perform better on the native [t’-a] [da] difference than do English infants aged 10 to 12 months. This finding was revealed on both an analysis of the proportion of subjects reaching criterion and in an analysis of A’ scores. Furthermore, the A’ scores were significantly better than chance for younger but not older infants. This evidence is consistent with previous cross-language research showing a developmental change in phonetic perception during the first year of life, but also increases our understanding of the mechanisms explaining age-related changes in speech perception. First, the present results show that the sensitivities of 6- to 8-month-old infants are indeed broad based. Second, older infants’ failure to discriminate [t’-a] vs [da] shows that the developmental reorganization in speech perception occurring at 10 to 12 months is not a function of simple linguistic exposure per se. Instead, the results are consistent with the notion that the shift is based on the phonemic status of the contrast.

This experiment also revealed that adults show moderate discrimination of the native language phonetic difference [t’-a] vs [da]. Adults did not perform as well as would be expected if tested with a native phonemic contrast. Nevertheless, adults discriminated better than chance on the present distinction unlike their typical chance level performance in this procedure with non-native contrasts. These results indicate that passive experience goes some way toward facilitating adult discrimination but is not sufficient to allow discrimination at the levels shown for meaningful contrasts.

IV. GENERAL DISCUSSION

The purpose of these studies was to further elucidate our understanding of the developmental reorganization in speech perception by investigating adult and infant discrimination of a phonetic difference that occurs in the native language but is not used to contrast meaning. Three experiments were designed to test two alternative hypotheses for explaining the functional reorganization. According to the first hypothesis, this reorganization occurs as a result of simply being exposed to phonetic variation in the language. If this hypothesis is correct, all age groups should have discriminated the difference between [da] and [t’-a] because this phonetic difference occurs in the linguistic input. According to the second hypothesis, the reorganization occurs as a function of the phonemic status of the phones in question. If this hypothesis is correct, only the younger infants, who as yet have no knowledge of the system of phonological contrasts in their native language, should have discriminated [t’-a] and [da], and both older infants and adults should have failed to discriminate this pairing.
In experiment 3, a significant number of English-learning infants aged 6 to 8 months discriminated \([t\text{-}a]\) from \([\text{da}]\) whereas English-learning infants aged 10 to 12 months did not. This pattern was evident in both the analysis of the proportion of subject at each age reaching criterion on the distinction as well as in the analysis of \(A'\) scores. Although discrimination by younger infants does not allow us to differentiate between the two hypotheses, it does provide additional understanding of younger infants' speech perception abilities. Discrimination by the infants aged 6 to 8 months shows that initial sensitivity to phonetic differences is very broad based, and possibly extends beyond sensitivity to just those phonetic differences that are used by one of the world's languages to contrast meaning.

In contrast to younger infants, older infants, who discriminated a meaningful English contrast \([t\text{-}a]\) vs \([\text{da}]\), failed to discriminate \([t\text{-}a]\) from \([\text{da}]\). Ten- to 12-month-old infants' failure to discriminate \([t\text{-}a]\) from \([\text{da}]\) shows that simple exposure alone is not sufficient to maintain discriminability at the end of the first year of life. These results are consistent with the second hypothesis that was posed: it is the match between the input and the phonological status of the distinction in question that governs infant speech perception by the end of the first year of life.

Data regarding adults did not provide clear support for either hypothesis. Experiment 1 showed that adult English speakers judge all exemplars of both \([t\text{-}a]\) and \([\text{da}]\) categories to be better examples of a written ‘‘D’’ than an ‘‘L.’’ Also, English adults considered exemplars of the syllable in the inappropriate context, \([t\text{-}a]\), to be equally good members of the ‘‘D’’ category as exemplars of the syllable in the appropriate context, \([\text{da}]\). Thus, the evidence from the adults in experiment 1, like that from older infants in experiment 3, appears to be consistent with the phonemic status hypothesis because there was no evidence that English adults can judge quality differences between \([t\text{-}a]\) and \([\text{da}]\) syllables. Nevertheless, when English adults were tested in an AX task with 500-ms ISI (experiment 2), they discriminated \([t\text{-}a]\) vs \([\text{da}]\) better than chance in three of four conditions. Even when tested in a category change procedure (experiment 3), adults discriminated this English phonetic difference at levels better than chance. This evidence appears to contradict the phonemic status hypothesis: English adults are able to detect the acoustic/phonetic differences between \([t\text{-}a]\) and \([\text{da}]\). Importantly, however, adult levels of performance were consistently below those shown for discrimination of native phonemic contrasts. These findings suggest that it is not essential for two phones to be used to contrast meaning for adults to be able to discriminate them, but that phonological status does influence the level of discriminability obtained.

Taken together, the results from this series of studies are most consistent with the hypothesis that perception of phonetic information by older infants and adults is best explained by the phonological status of the distinction in question. Together with previous work (Best, 1994; Best et al., 1995, 1988; Polka, 1995; Werker, 1989, 1994; Werker and Pegg, 1992; Werker and Tees, 1984a) these results provide strong support for the notion that it is the map between the phonetic input and the phonological system that best explains age-related shifts in perception of fine phonetic detail. In this particular set of experiments, support is gained for the notion that it is the set of phonological contrasts that provides the strongest influence on age-related changes in phonetic perception.

Two caveats need to be kept in mind when considering this conclusion. First, although adults were unable to distinguish \([\text{da}]\) and \([t\text{-}a]\) in experiment 1, and performed more poorly on this distinction than is typically shown for native contrasts in experiments 2 and 3, their level of performance was still significantly better than chance in three of four conditions in experiment 2 and in both orders in experiment 3. As well, the performance of the adults in experiment 3 was significantly better than that of 10- to 12-month-old infants. Thus it is incorrect to conclude that phonological status entirely accounts for sensitivity to fine phonetic detail in adults, and it is essential to offer some account of how performance improves again between late infancy and adulthood. We think the best explanation for these results is that although a phonemic (meaning based) listening strategy is the most robust listening style shown by adults, the ability to redirect attention to other acoustic/phonetic differences is also present. This view is consistent with evidence provided by numerous studies on the adaptability of adults’ listening strategies and acquisition of second languages (e.g., Pisoni and Tash, 1974; Werker, 1994; Werker and Logan, 1985; see Strange, 1995 for a review). Strategic processing may account for the superior performance of adults in comparison to 10- to 12-month-old infants: older infants are unable to deploy selective listening strategies and are more constrained to listen to speech for that information that is most meaningful for them.

The second caveat concerns the nature of ‘‘meaning’’ in infants of 10 to 12 months of age. By meaning, we intend the notion of functionally useful. Although infants of 10 to 12 months undoubtedly comprehend some words (Fenson et al., 1994), they are speaking few words at this age. Furthermore, there is little evidence that they are, at this age, able to distinguish between phonetically similar words (Stager and Werker, 1995, 1997; Werker and Pegg, 1992). Thus it may be that what infants are listening for at this time is ‘‘possible words’’ rather than actual words that map sound on to meaning. This is the functional aspect of infants’ listening strategies.

There is considerable evidence that by 10 months of age, but not before, infants have sensitivity to the stress patterns (Jusczyk et al., 1993a), acceptable (Jusczyk et al., 1993b) and common (Jusczyk et al., 1994) phonotactic sequences, and other distributional properties of the native language (Aslin et al., 1996). They have also the ability to coordinate two sources of information (Lalonde and Werker, 1995) allowing simultaneous consideration of stress pattern and position-specific phonetic information (Morgan and Saffran, 1995). With the emergence of these abilities, infants are able to extract even unfamiliar word forms (Meyers et al., 1996) and recognize words in fluent speech (Jusczyk and Aslin, 1995; Jusczyk et al., 1995). Listening for possible words without actually mapping those words on to specific objects and events would yield the same outcome for the older in-
fants. It would bias them to treat the syllables \([t^\text{a}a]\) and \([\text{d}a]\) as equivalent. Since unaspirated \([t^\text{a}]\) does not occur in syllable initial position in English, both \([t^\text{a}a]\) and \([\text{d}a]\) could be equivalent English word forms.

In summary, this series of experiments has provided convincing support for the hypothesis that it is the phonological status of phonetic detail that best explains the age-related changes in speech perception across the first year of life. Although adult speech perception is also influenced by phonological status, adults seem to be able to apply flexible listening strategies when required. With these results in hand, future research can focus directly on the question of what “phonological status” entails to the infant of 10 to 12 months, on the conditions under which adults can be sensitized to nonphonemic phonetic variation, and on the age at which adultlike flexible strategies emerge.

ACKNOWLEDGMENTS

This work was supported in part by a Post Graduate Scholarship from the Natural Sciences and Engineering Council of Canada (NSERC) and a Studentship from the Medical Research Council of Canada (MRC), awarded to the first author as well as a research grant from NSERC awarded to the second author (OGP0001103). We would like to thank the adults, parents, and infants who participated in this research. Finally, we would like to acknowledge the efforts of several individuals who participated in booking appointments, assisted in testing infants and adults, and/or generally kept the laboratory running smoothly: Diane Hobday, Gary MacIsaac, Brian Moorehead, Tracey Smiley, and Carmen Swanson.

1One person had been briefly exposed to Greek at 4 years of age.
2Prevoicing is common in spoken English. We sampled several speakers before finding one male who did not prevoive or devoice \(\text{id}/\).
3We did not include a condition with “TA” page heading because the set of syllables initially position in English, both \([t^\text{a}a]\) and \([\text{d}a]\) could be equivalent English word forms.
4The formula for \(A'\) is \(0.5 \times (H - FA) / (1 + H - FA) / 4H(1 - FA)\), where \(H\) is proportion of hits and \(FA\) is proportion of false alarms (from Grier, 1971).
5It is not unusual in this procedure to have high attrition rates, but it is possible that \([\text{pra}]\) vs \([\text{da}\]) may be somewhat more difficult for infants than other phonemic contrasts on which they have been tested.
6Even though very low \(A'\) scores are invalid, all the analyses reported below were repeated including the three low \(A'\) scores and every significant result reported remained significant.
