Young Infants' Pitch Generalization over Unknown Words
Amritha Mallikarjun, Spring SURG Proposal, Carnegie Mellon University

Abstract
Studies have shown that young infants who are given words in a certain affect or pitch cannot recall the word if it is presented in a different affect or pitch to them (Houston & Jusczyk, 2000, Singh et al., 2004). This makes word learning difficult, as infants cannot understand the same word spoken by one person when produced by a different speaker. Infants have been shown, however, to generalize words over multiple different affects if given words in happy, sad, angry and fearful tones (Singh, 2008). This study expands on Singh’s work in two ways. First, it assesses how varying the pitch of the input given to 7.5 month olds can help them generalize words given at a higher pitch to words at a lower pitch. Second, it extends these questions about generalization to input involving fluent speech, as opposed to the words presented in isolation in prior research. This is an important question because most words are presented to infants in fluent speech rather than in isolation. The single pitch condition features a steady high-pitched language followed by test words in a low pitch. The varying pitch condition features an interleaved high pitch and medium pitch language followed by the same low pitch test words. This will help us further understand how infants can use the variation of input they receive to learn to generalize words.

Research Question and Significance
After my previous SURF produced interesting results suggesting that infants can use an accent cue to tell two different languages apart, I realized that even when people speak the same language, the pitch, tone, affect, amplitude, and many other characteristics change between speakers. It is remarkable how infants are soon able to generalize the idea that regardless of who might say the word “cup” to them, whether they have a high-pitched, fast-paced female voice or a booming, slow, mid-range male voice, the label refers to the same kind of object.

When young infants begin the word learning process, their memory representations of words include indexical information about the speaker, like the auditory and emotional characteristics of the original speaker’s voice. This is likely helpful, as it helps infants to discover which acoustic features are relevant in their native language. For example, it would be crucial for a Chinese infant to consider pitch of speech when they listen to adults talk, but it is not as important for an American infant.

By storing detailed acoustic information, infants can eventually discover which acoustic features are relevant in their own language. However, this ‘excessive’ detail means that younger infants generalize poorly over different voice qualities. After listening to the word “cup” spoken by a female, 7.5 month-olds do not recognize “cup” produced by a male speaker as being the same word (Houston & Jusczyk, 2000). Singh, Morgan & White (2004) also showed that infants familiarized with words in a happy voice do not listen longer to the words if it is presented in a neutral voice. This characteristic is not conducive to word learning, as it shows that young infants regard, for example, the word “ball” in a female voice to be a distinct word from “ball” in a male voice.

As infants progress in age, they discover which acoustic cues are relevant in their native language, which helps them to disregard cues like gender and amplitude so they can generalize that whenever someone says “ball”, it is the same word (e.g., Graf Estes, 2012). One feature of the input that helps infants discover which acoustic features are relevant is variability: those acoustic features that are not relevant to the identity of a word should vary more widely (across instances of that word) than features that are
Young Infants’ Pitch Generalization over Unknown Words
Amritha Mallikarjun, Spring SURG Proposal, Carnegie Mellon University

relevant (such as its phonemic components). Consistent with this, variability in the input appears to influence infants’ generalization. Singh (2008) has shown that presenting 7.5-month-old infants with happy, sad, angry, and fearful versions of different words helps them to recognize the words when they are presented with happy or neutral words in an unfamiliar sentence. Infants that are only given happy words, on the other hand, cannot generalize to the neutral words within the sentences.

My experiment will look at these generalizations in 7.5 month olds in the context of statistical learning, which is another method infants use to learn words. For example, a baby hears the word “pretty” very often in phrases like “pretty shoes”, “pretty baby”, and “pretty eyes”. This leads the baby to understand that the syllables “pre” and “ty” occur together, and together most likely comprise one word. In this manner, babies learn words as they figure out which syllables occur frequently next to one another when they hear people talk (Thiessen & Saffran, 2003). I will also run the experiment on adults, to make sure that the stimuli is at an appropriate difficulty level for the infants.

This experiment will show lexical generalization in the more natural context of fluent speech and will help to demonstrate that infants can harness variability to help learn words more efficiently. If this experiment indicates that varied input given in normal, unbroken speech helps infants generalize, it will show that infants can benefit from listening to many different speakers as they grow up to help expose them to the multiple different pitches, affects and voice qualities that they can use to help them decipher their own language.

**Project Design and Feasibility**

This experiment requires an artificially synthesized (no human voice) language stream and will rely on habituation to see if the infants have correctly separated the languages. This experiment shall use existing languages created by Lew-Williams and Saffran in their 2012 statistical learning experiment. The language contains 4 disyllabic (CV.CV) words, as seen in the included word table. I will synthesize three versions of the language for testing. The first version will be created in a medium pitch, around normal female speaking tone of 200 Hz. The second version will be created in a pitch 20 Hz higher, and the third will be at a pitch 20 Hz lower. The high and medium pitch versions will be used during the initial exposure phase, and the low pitch version will be used for testing.

There will be an exposure phase at the start of this experiment where the infants listen to 4 minutes of either only the high pitched language or an interleaved stream of the high- and medium-pitched versions, and then a testing phase in which the infants listen to either a word in the low-pitched version (baku, tiro, pido, or lagu) or a part-word, made with the last syllable of one word and the first syllable of the next (kuti, ropi). This would be the equivalent of “ty-ba” in the pretty baby example: infants hear those syllables spoken together, yet they are not a word.

The words in each language stream will be structured in a way that in the test condition, the infants must combine statistics between the two languages in order to successfully learn the words. To achieve this, in the high pitch language the word *tiro* will always follow the word *baku*, while in the medium pitch language the *tiro* can come after any word. If the infants do not combine statistics, learning the high pitched language...
Young Infants’ Pitch Generalization over Unknown Words  
Amritha Mallikarjun, Spring SURG Proposal, Carnegie Mellon University

would be the equivalent of always hearing “pretty-baby” together – it would be impossible to tell if it is just one large word (prettybaby) or two separate words. This way we can know if the infants successfully understand that the two pitch variations given to them in the exposure phase contain the same words, because infants would be confused about what constitutes a word and would not pass the test trials. The test trials are comprised of a set of gaze tests. I play a word or a part-word in the low pitch variation while the video of a bouncing ball is on the screen. I then time the amount that the baby looks at the screen for each trial. I predict that if the baby has successfully generalized the words from the high- and medium-pitch variations, they would gaze longer at the words and less long at the part-words, because they are more familiar with the whole words. If the babies think that the low-pitch variation is an entirely separate language, they would become confused, since they do not recognize any of the words. This would lead them to gaze an equal amount of time at the words and part-words, as they would not be able to tell them apart.

In the adult version of this experiment in EPrime, the exposure phase will be the same, but in the test trials, adults will get a pair of one word and one part-word, and they will be asked to choose which of the two is a word.

**Background**
I have been working in the Infant Language and Learning Lab with Dr. Thiessen for two years now, where I have experience recruiting parents to bring in their infants and running both infant and adult experiments. I am also taking Infant Language Development with Dr. Thiessen. I have worked many times with EPrime in Cognitive Research Methods and in lab. I have also previously done a SURF involving a similar kind of statistical learning experiment.

**Feedback and Evaluation**
I will meet with my advisor, Dr. Thiessen, weekly to check up on this project. I can also ask Ashley Episcopo, the Infant Language and Learning Lab manager, for quick help if I need it.

**Dissemination of Knowledge**
I will present at Meeting of the Minds with a poster, and I will most likely also present at lab meeting for the Infant Language and Learning Lab.

**Budget**

<table>
<thead>
<tr>
<th>Money Requested</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>$200</td>
<td>Testing adults to make sure that the language is at the appropriate difficulty for infants ($5 per participant, with 40 participants)</td>
</tr>
<tr>
<td>$120</td>
<td>Purchasing books (approx. $6 per book) to give as compensation for 20 families who bring infants in to the lab for the experiment, which would give me 10 infants per condition.</td>
</tr>
</tbody>
</table>