A 5

Can adult phonetic categories be predicted based on statistical distribution alone?

Megha Sundara
School of Communication Sciences & Disorders, McGill University, Montreal
E-mail: msunda@po-box.mcgill.ca

ABSTRACT

Recently there has been support for learning of phonetic categories based on sensitivity to distributional differences. Distribution-based approaches are powerful in that they make a priori predictions of discrimination performance for certain non-native contrasts. In the present set of studies we examined the limits of a distribution-based approach to predict phonetic categories of adults. We tested Canadian English, Canadian French, native Hindi and simultaneous bilingual Canadian English-Canadian French listeners on discrimination of a dental-alveolar contrast. These groups of listeners differ on whether they hear a unimodal or bimodal distribution of dental and alveolar stops during acquisition. We find that contrary to prediction, listeners exposed to a unimodal distribution show discrimination performance ranging from poor through intermediate to good. Listeners exposed to bimodal distribution, following prediction, show good discrimination performance. We discuss specific problems of applying distribution approaches to acquisition of categories in a natural context.

INTRODUCTION

Acquisition of phonetic categories is a complicated task for the infant. The input tokens vary a great deal on a large number of phonetic dimensions. Some of these dimensions are contrastive in the infant's native language while others are not. What complicates the task even more is that at no time is the infant explicitly aware of how many categories there are in the input language. The infant has to learn to recognize which dimension(s) to pay attention to and which to ignore.

One way an infant can learn which dimensions are relevant is by tracking clusters in the input. Phonetic tokens that form categories tend to cluster together, away from other clusters. Thus infants who detect more than one cluster learn to detect differences between these categories while infants who do not tend to ignore them. In other words, infants exposed to a bimodal distribution tend to discriminate tokens from these two distributions, while infants exposed to unimodal distribution on the same dimension do not.

There is substantial support for distribution based learning in perceptual acquisition of speech sounds. There is evidence that infants are sensitive to statistical distribution [1]; there is also evidence for better discrimination, in both adults [2] and infants [3], based on mere exposure to bimodal distribution of input. These studies have been remarkable in that they often demonstrate robust results using single-talker stimuli in one or two vowel contexts with exposure restricted to a few minutes in controlled laboratory settings. Differently stated, it appears that a distribution-based learning is a viable mechanism for acquisition of phonetic categories.

Given that a statistical distribution based mechanism is so powerful, it is important to investigate its limits. In language acquisition, adults typically have extensive experience with multiple speakers producing tokens in diverse vowel contexts. Phonetic category learning in this context necessitates treating members of the same category as equivalent in addition to learning to discriminate between two tokens from different categories. Finally, statistical distribution and functional load (or phonemic status) are almost always confounded, implying that sounds that are distributed bimodally also contrast meaning. In view of these differences between previous investigations of distribution-based learning in controlled settings and natural language acquisition, there is a need to see how far the statistical regularities in the input can take us in predicting categories subsequent to language acquisition in a natural context.

In the experiments presented we tested how 4 groups of listeners differing in language exposure during acquisition discriminate a voiced dental-alveolar stop contrast. The dental-alveolar place contrast for stops is highly marked. It is rare across the inventories of the world's languages (in O'odham and a few Australian aboriginal languages), and even when present has highly restricted phonotactics (in Malayalam).

The 4 groups of listeners differ systematically on whether they hear a unimodal or bimodal distribution of dental and alveolar stops in their native language input. We tested two predictions based on statistical distribution differences. First, adult listeners exposed to unimodal distributions are poor at discriminating between dental and alveolar categories. Second, adult listeners exposed to bimodal distributions, in the absence of phonemic status for the contrast, are good at discriminating between dental and alveolar categories.

ISBN 1-876346-48-5

EXPERIMENT I: UNIMODAL EXPOSURE & CATEGORICAL DISCRIMINATION

In this experiment, we tested 3 groups of adults; monolingual Canadian French (CF), monolingual Canadian English (CE) and native Hindi listeners. In their native language, CF listeners hear dental stops in all positions [4]. CE listeners hear voiced alveolar stops in all positions, except in cases where they are produced before inter-dental fricatives. In the latter case, stops may be produced in a dental rather than an alveolar position. Both groups hear a unimodal distribution; on a discrimination task for the dental-alveolar contrast we expect them to perform poorly. Native Hindi listeners hear dental stops in Hindi; on a discrimination task for the dental-alveolar contrast we expect them to perform poorly.

SUBJECTS

We tested 8 adult of monolingual CF, monolingual CE and native Hindi listeners. Subjects had no history of speech, language or hearing impairment. Their language background was assessed using a detailed language questionnaire.

To be included in the monolingual CF (or CE) group subjects had to meet the following four criteria. First, both parents of the subjects' were monolingual speakers of CE (or CF). Second, their schooling was completed in CE (or CF). Third, subjects rated their ability in their native language with a minimum of 6 on a scale of 1-7 (where 7 represents native-like ability while 1 represents no ability at all). If they had any knowledge of the non-native language[i] they rated it below 3 on the same scale. A fluently bilingual interviewer confirmed their self-ratings of level of proficiency in the non-native language. Four, they had spent no time in a country where a language other than their native language was spoken.

To be included in the native Hindi group, the subjects had to meet the following criteria. First, both parents of subject's were native Hindi speakers. Second, they used Hindi frequently and consistently in conversation with friends and family. Third, subjects rated their ability in Hindi with a minimum of 6 on a scale of 1-7 (where 7 represents native-like ability while 1 represents no ability at all). A native speaker of Hindi confirmed their self-ratings of level of proficiency in conversational Hindi. Four, they had arrived in Canada less than 1 year ago.

STIMULI & METHODS

We recorded 3 male monolingual CE and 3 male monolingual CF talkers producing bisyllabic words with voiced coronal stops in syllable initial position. Criteria for selection of monolingual talkers were as described above for monolingual listeners. Subsequently, we excised CV syllables from these bisyllabic words for use as stimuli.

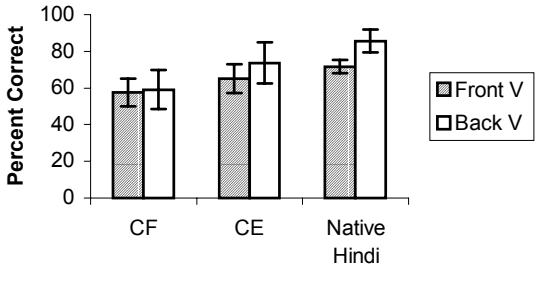
Excised syllables were digitized (at 22KHz and 16 bit) and analyzed using PRAAT. Natural tokens were subsequently edited to equalize the VOT across the two language tokens. CF & CE differ systematically in the VOT used to signal voicing. Voiced consonants in CF have lead VOT or prevoicing while voiced consonants in CE have short lag VOT. For all CF tokens the prevoicing was edited out. This made the VOT for the two sets of tokens comparable. Tokens in the final set were selected such that the distribution of fundamental frequency, amplitude and duration of the syllables overlapped completely in the two languages. Front vowel /æ/ and back vowels /o/, /ɔ/ and /ɑ/ were used. These vowels are acoustically similar in the two languages.

We assessed perception performance with a categorical discrimination task (AXB) using BLISS. Each trial was made up of three syllables, each syllable produced by a different talker. Testing was done in two blocks. Each block consisted of 288 trials with an inter-stimulus-interval of 1500msec and an inter-trial interval of 4000msec. In blocks I, tokens were presented in front vowel context. In block II, tokens were presented in back vowel context. Order of testing for the blocks was fixed. First, the subjects were presented with a 10 trial practice session; then they completed block I with the front vowel; followed by block II with the back vowel. Subjects were given a break in the middle of each block. Accuracy and reaction times were measured. In this study we report accuracy data only.

RESULTS & DISCUSSION

Percent correct responses of the three groups across the two vowel conditions are presented in Figure 1. The pattern of performance for the front and back vowel condition was similar across the groups. However, the overall performance on the AXB task was poorer in the front vowel than in the back vowel context. We compared percent correct performance on the front and back vowel contexts separately using Kruskal Wallis test.

Figure 1: Listeners with unimodal exposure



There was a significant difference between the 3 groups on the front ($p<0.01$) as well as the back vowel ($p<0.01$) context. CF listeners showed poor discrimination performance, 57% for the front vowel and 59% for the back

vowel context. CE listeners showed intermediate level of discrimination performance, 65% for the front vowel and 73% for the back vowel context. Native Hindi listeners showed good discrimination performance, 71% for the front vowel and 85% for the back vowel context.

An approach based on statistical distribution predicts that groups that are exposed to unimodal distributions for the contrast are poor at discriminating it. This prediction was supported for only 1 group of the 3 we tested. Only CF listeners show poor performance on the dental-alveolar contrast. CE and native Hindi showed intermediate and good performance on this contrast.

One explanation why CE did not show a poor performance could be that these listeners are exposed to the dental voiced stops (though in a highly restricted context) as allophones. In addition, English listeners are also exposed to a dental-alveolar place distinction for fricatives [5]. Thus, perhaps the input for the CE listeners is not accurately described as unimodal. We encounter a similar problem in categorizing the input pattern for native Hindi as unimodal or bimodal. While native Hindi speakers are exposed to a unimodal distribution for the dental-alveolar contrast; they are exposed to a bimodal distribution for coronal place distinction (dental as well as retroflex place) for stops. Thus, based on distribution along coronal place, as opposed to just the contrast in question, we would get a more graded prediction for the CE and the native Hindi listeners.

In summary, listeners exposed to unimodal distributions are not good at discriminating the dental-alveolar contrast. Based simply on statistical distribution we are unable to predict the level of discrimination performance for listeners hearing a unimodal input. There is a need to empirically determine dimensions over which distributions are calculated. This is pertinent specifically to any discussion of acquisition of categories in a setting outside of a laboratory.

EXPERIMENT II: BIMODAL EXPOSURE AND CATEGORICAL DISCRIMINATION

In this experiment we tested a group of simultaneous bilingual CF-CE listeners. These subjects hear dental stops in CF and alveolar stops in CE. They are systematically exposed to bimodal distribution of the dental-alveolar contrast across their two native languages; this place distinction is not phonemic (non-functional) in either one of their native languages. We expect them to perform very well in a discrimination task for the dental-alveolar contrast.

SUBJECTS

We tested 6 simultaneous bilingual CF-CE listeners. Subjects had no history of speech, language or hearing impairment. Their language background was assessed using a detailed language questionnaire.

To be included in the early simultaneous bilingual group subjects had to meet the following four criteria. First, subjects had learnt both CE & CF simultaneously at home from parents, each of whom was a native speaker of one of them[ii]. Second, their schooling was completed either in bilingual schools, or at different points in CE and CF. Third, subjects rated their knowledge of both languages with a minimum of 6 on a scale of 1-7 (where 7 represents native like ability while 1 represents no ability at all). A fluently bilingual interviewer confirmed their self-ratings of level of proficiency in both languages. Four, they were using both languages consistently within the home and the work context.

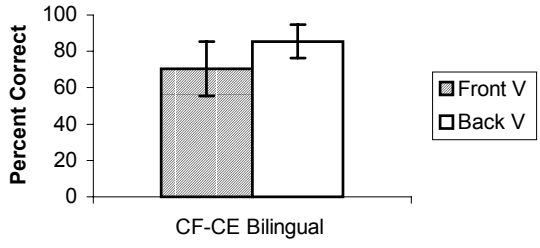
STIMULI & METHODS

Same as experiment I.

RESULTS

Percent correct responses of the bilingual group across the two vowel conditions are presented in Figure 2. The overall performance on the AXB task was poorer in the front vowel than in the back vowel context. Bilingual CF-CE listeners showed good discrimination performance, 70% for the front vowel and 85% for the back vowel context. We compared the performance of bilingual listeners to that of native Hindi listeners in Experiment I and found no significant difference between the two.

Figure 2: Listeners with bimodal exposure



An approach based on statistical distribution predicts that groups that are exposed to bimodal distributions for the contrast are good at discriminating it. This prediction was supported for the bilingual group, even when this distinction is not phonemic (meaningful) in either language of the listeners.

GENERAL DISCUSSION

Statistical-distribution based approaches are a strong tool to make a priori predictions about discrimination performance. This is particularly important in cases like predicting phonetic categories of bilingual listeners. Using the

 ISBN 1-876346-48-5

statistical distribution approach minimizes the need to make assumptions about underlying organization of bilingual language systems. The experiments in this paper were designed to test how well we can predict adult phonetic categories based on the distribution of the input alone.

Data from experiments I & II offer us insights into the strengths and weaknesses of a statistical distribution approach. We present data to show that CF listeners who are exposed to a unimodal distribution discriminate the dental-alveolar contrast poorly. We also present data that discrimination performance for all listeners exposed to unimodal distributions is not poor. CE listeners as well as native Hindi listeners consistently perform better than CF listeners. Thus, the statistical distribution approach fails to predict the level of performance in different groups of listeners exposed to a unimodal distribution.

As mentioned in the discussion for experiment I, one could make a more graded prediction by redefining the dimension on which the decisions for unimodal or bimodal distributions are made. This would fit the data better. However, one advantage of the statistical approach is that it allows one to make a priori predictions and redefining the dimension on a case-by-case basis dilutes that. We need more empirical investigation to determine the optimal dimension on which to calculate distributional differences such that accurate predictions can be made for prediction of discrimination performance for natural categories.

We also present data that bilingual listeners exposed to a bimodal distribution discriminate the dental-alveolar contrast very well. Thus, it is accurate to predict that adults exposed to bimodal distributions develop good discrimination for the contrasts in question. Finally, as adult listeners are able to use top-down influences of a fully developed phonology in phonetic discrimination, distribution differences might be most useful in predicting infant categories in the absence of fully developed phonology. We are currently testing 10 to 12-month-olds to further investigate the role of distribution differences on infant phonetic categories.

REFERENCES

[1] P W Juscyzk, P Luce, J Charles-Luce, "Infants' sensitivity to phonotactics patterns in the native language," *Journal of Memory and Languag*e, vol. 33, 630-645, 1994.

[2] J Maye and L Gerken, "Learning phonemes without minimal pairs," in *Proceedings of 24th BUCLD*, Somerville MA: Cascadilla Press, pp. 522-533, 2000.

[3] J Maye, J Werker and L Gerken, "Infant sensitivity to distributional information can affect phonetic discrimination," *Cognition*, vol. 82, B101-B111, 2002.

[4] M Picard, "*Phonetics and phonology for ESL and TESL teachers*," Montreal, Canada: Concordia University Press, 2001.

[5] L Polka, "Cross language speech perception in adults: Phonemic, phonetic, and acoustic contributions," *Journal of acoustic society of America*, vol. 89, 2961-2977, 1991.

[i] Most people educated in Canada receive formal instruction in both languages at school; this instruction is mainly in reading and writing with no emphasis on speaking or listening skills. Thus, proficiency of the students in the second language varies vastly.

[ii] A large numbers of families in Canada have been following a one parent-one language tradition while bringing up children making this kind of a home environment fairly standard among inter-language marriages.

ISBN 1-876346-48-5