The Limits of Multidimensional Category Learning

Martijn Goudbeek [1,2], *Daniel Swingley,* [1,3] *& Keith R. Kluender* [4]

[1] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[2] Geneva Emotion Research Group, University of Geneva, Geneva, Switzerland
[3] Department of Psychology, University of Pennsylvania, Philadelphia, USA
[4] Department of Psychology, University of Wisconsin, Madison, USA

goudbeek@pse.unige.ch, swingley@psych.upenn.edu, krkluend@wisc.edu

Abstract

Distributional learning is almost certainly involved in the human acquisition of phonetic categories. Because speech is inherently a multidimensional signal, learning phonetic categories entails multidimensional learning. Yet previous studies of auditory category learning have shown poor maintenance of learned multidimensional categories. Two experiments explored ways to improve maintenance: by increasing the costs associated with applying a unidimensional strategy; by providing additional information about the category structures; and by giving explicit instructions on how to categorize. Only with explicit instructions were categorization strategies maintained in a maintenance phase without supervision or distributional information.

Index Terms: auditory categories, distributional learning, multidimensional categorization, supervised learning.

1. Introduction

Both infants learning a first language and adults learning a second language are faced with a daunting multidimensional categorization problem: the phonetic categories of their language differ on a large number of relevant dimensions. A famous report by Lisker, for example, indicates at least sixteen relevant cues to distinguishing the voiced and unvoiced plosives in the utterances /rapid/ and /rabid/ [1]. Distributional learning is a reasonable candidate for being a domain general category learning mechanism. Furthermore, it is undoubtedly available even to infants, given that they begin to learn phonetic categories in their first year [2], [3].

Rapid distributional learning of visual and auditory categories has been demonstrated in laboratory settings [4], [5]. In [4], infants were presented with stimuli varying on an artificial voice-onset-time continuum ranging from [da] to [ta]. Following exposure to a unimodal or bimodal stimulus distribution, infants listened to stimulus sets that were either alternating or non-alternating. Only infants that were exposed to the stimuli drawn from the bimodal distribution preferred alternating to non-alternating stimuli. A similar sensitivity to distributional cues was found for adults [6].

Since speech is an inherently multidimensional signal containing many different regularities, a viable account of distributional learning of speech categories must incorporate learning of categories that are defined over more than one dimension. Learning of multidimensionally varying category structures has mostly been studied in adults learning visual categories [7]. These studies show that, at least in the visual domain, participants can be brought to entertain a multidimensional categorization strategy, but only with the aid of trial-by-trial feedback [8].

Figure 1 shows the ineffectiveness of unidimensional categorization rules in multidimensional categorization problems. The characters in the plot correspond to stimuli varying in two dimensions. These cluster in two categories (the ellipses) that are diagonally oriented. This diagonal orientation means that effective categorization requires consideration of each dimension. Each point is plotted as either correctly categorized in category A (the circles) or category B (the pluses) or incorrectly classified and consequently labeled as an error (the asterisks). Perfect multidimensional categorization would be revealed by each ellipse being plotted with a single character, as displayed in the right panel. In the left panel only dimension 1 is being used (all stimuli left of the line are categorized as category A and all stimuli right of the line as category B), leading to many errors. Given the costs of such a unidimensional strategy, listeners should shift to a multidimensional strategy in order to avoid making many errors.
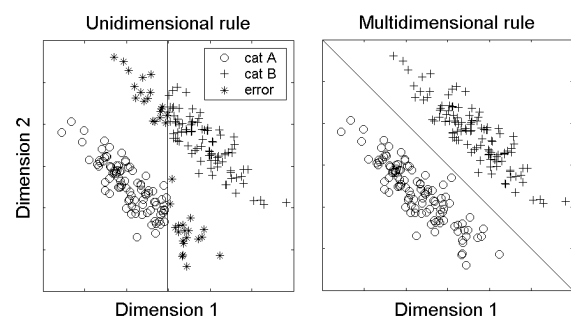

Figure 1: *Correct and incorrect categorizations with uni- and multidimensional categorization rules.*

Earlier experiments investigating multidimensional auditory category learning have shown poor maintenance of multidimensional category learning [9]. These experiments consisted of a learning phase and a maintenance phase. In the learning phase, the stimuli contained distributional information (see the left panel of Figure 2), whereas the stimuli in the maintenance phase did not contain any distributional information (see the right panel of Figure 2). In the maintenance phase listeners usually revert to a unidimensional strategy. Since there is no distributional

information in the maintenance phase favoring the use of one dimension over the other, this inability to maintain the previously learned multidimensional strategy is puzzling.

The stimuli in our experiments were constructed by defining a multidimensional stimulus space spanned by the dimensions *formant frequency* (measured in ERB) and *duration* (measured in DUR) [10]. These dimensions were approximately equalized in terms of their just noticeable differences. Stimulus tokens were inharmonic tone complexes varying in length, defined by the dimension duration and filtered at a single resonance peak, defined by the dimension formant frequency. All stimuli were RMS matched to ensure a constant sound pressure level of 65 dB.
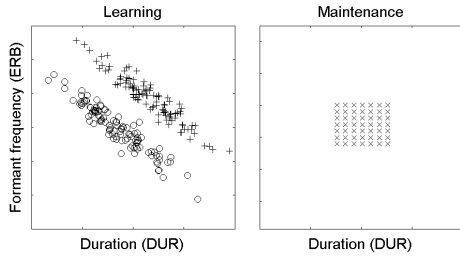

Figure 2: *Distributional characteristics of the learning and maintenance phase.*

The experiments presented here attempted to improve multidimensional categorization performance in the maintenance phase in two ways. In Experiment 1 the distributional properties of the categories in the learning phase were manipulated to increase the costs of a unidimensional strategy. In Experiment 2 feedback and additional sources of information were added to the learning phase to stress the multidimensional nature of the categorization problem to the listeners.

2. Experiment 1

In Experiment 1 the distance between the category means and their standard deviations were manipulated to increase the incentive for a multidimensional categorization strategy in the learning phase. Figure 3 shows the distributions of the three conditions and Table 1 lists the proportion between the mean and standard deviations relative to those in Condition 1 and the resulting optimal percentages correct for a unidimensional or multidimensional strategy.

Table 1: *Proportions for distance between the means and sd's with associated optimal percentage correct.*

Cond.	Proportion μ	Proportion σ	Uni	Multi
1	1	1	82%	100%
2	0.66	0.66	70%	100%
3	0.25	0.25	57%	100%

Table 1 shows that the extent to which multidimensional categorization was beneficial depended on the distributional properties. A smaller distance between the means and a smaller standard deviation lead to more categorization errors. However, these categories might also be more difficult to separate due to noise in perception and memory.
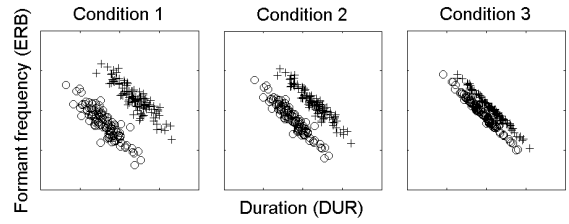

Figure 3: *Category structures of the three conditions of Experiment 1.*

All 30 participants were psychology students from the University of Wisconsin. The number of participants was 12, 11 and 7 in Conditions 1, 2, and 3 respectively. They were instructed to assign sounds to one of two buttons and received trial-by-trial feedback on each categorization. Each stimulus was presented in two randomized blocks, resulting in 448 trials (112 stimuli x 2 categories x 2 blocks). In condition 3 (the most difficult condition), listeners received an additional learning block, resulting in 672 (112 x 2 x 3) trials.

After the learning phase, listeners entered the maintenance phase and categorized the maintenance stimuli as they saw fit. In the maintenance phase, feedback as well as distributional information were absent.

2.1. Results

The results were analyzed using the β-weights yielded by logistic regression. In summarizing participants' performance, simply computing mean beta weights over subjects could mischaracterize the results, e.g., producing an illusion of multidimensional categorization by a set of listeners, when in fact half used one dimension and half used the other.

To avoid this problem, we derived a measure that integrates performance on both dimensions by performing a polar transformation on the β-weights where the x axis represents duration, and the y axis formant frequency [9]. Consider the line connecting the origin with a point (β_{dur}, β_{freq}). The angle Φ with the horizontal axis in radians indicates the contribution of each dimension. Here, Φ ranges between π and -π radians. When Φ equals 0.5π (a vertical line with a β_{freq} of 0) , listeners only use formant frequency, when Φ equals 0 (a horizontal line with a β_{dur} of 0), listeners use only duration, and when Φ is 0.25π listeners use both dimensions equally. Mathematically, Φ is defined as follows:

$$\phi = \arctan(\beta_{freq}/\beta_{dur}) \; if \; \beta_{dur} \leq 0 \qquad (1a)$$

$$\phi = \arctan(\beta_{freq}/\beta_{dur}) + \pi \; if \; \beta_{dur} < 0; \; \phi - 2\pi \; if \; \phi > \pi \quad (1b)$$

The length A of this line is related to the size of the β-weights and indicates the consistency of the categorization. Large values of A indicate that listeners agree with each other, indicating consistency across subjects (note that if listeners are all wrong in the same way, A will have a large value; it measures consistency, not accuracy). Mathematically A is defined as follows:

$$A = \sqrt{(\beta_{dur}^2 + \beta_{freq}^2)} \qquad (2)$$

The results shown in Table 2 indicate that performance in the learning phases of Conditions 1 and 2 was good. The mean value of Φ was close to 0.25π indicating the use of both dimensions and A was large indicating consistency across participants. This multidimensional strategy was lost in the maintenance phase, with Φ's close to zero (indicating the often observed preference for duration). The large A indicated consistent use of a unidimensional (wrong) strategy. The small A and the small number of listeners using both dimensions in Condition 3, suggest that the close spacing of the two categories reduced overall performance, and also reduced the tendency toward multidimensional categorization. The larger A in the maintenance phase indicates listeners' sensitivity to the (absence of) distributional information and feedback, as they were better able to maintain an (incorrect) unidimensional categorization strategy.

Table 2: *Mean Φ (in π) and A values and the number of listeners using both dimensions in Experiment 1.*

Cond.	N		Learning	N_{multi}	Maintenance	N_{multi}
1	12	Φ (σ)	0.27 (0.03)	12	0.12 (0.20)	3
		A (σ)	1.44 (0.57)		1.27 (0.41)	
2	11	Φ (σ)	0.29 (0.04)	11	0.05 (0.20)	3
		A (σ)	1.24 (0.57)		1.06 (0.79)	
3	7	Φ (σ)	0.11 (0.29)	4	0.00 (0.26)	0
		A (σ)	0.29 (0.23)		0.79 (0.63)	

Statistical evaluation of Φ confirmed that listeners were unable to maintain a multidimensional categorization rule. In the learning phases of Conditions 1 and 2, Φ differed significantly from both 0 and 0.5π (t_{min} = 18.1, p < 0.05) indicating multidimensional categorization. In the maintenance phase Φ did not differ significantly from 0 but did vary significantly from 0.5π (t_{min} = 6.4, p < 0.05), showing the overall preference for a unidimensional categorization rule with duration as the relevant dimension.

After ensuring all A's differed significantly from zero in all conditions (t_{min} = 3.3, p < 0.05) an ANOVA with A as a dependent measure and condition as the independent variable was conducted. This showed the conditions to differ significantly in the learning phase (F [2,27] = 11.7, p < 0.05). but not in the maintenance phase (F [2,27] = 1.36, n.s.). Post hoc tests (Tukey HSD) on the learning phase showed performance in Condition 3 to be significantly worse compared to the other two.

2.2. Discussion

The goal of Experiment 1 was to increase the number of listeners that used a multidimensional categorization rule in the maintenance phase by manipulating the distributional properties of the categories in the learning phase. In effect, the condition that punished unidimensional categorization the most, Condition 3, had the worst multidimensional performance. Apparently, listeners did not respond to these properties. An explanation for this result might be that the categories in Condition 3 were too difficult to separate for the listeners. The distance between the means may have simple

been too small to be perceptually separable. Another possible explanation for this finding is that there was so much negative reinforcement (the error rate for an initial unidimensional strategy is very high) that listeners gave up before they could discover the multidimensional rule.

3. Experiment 2

The goal of Experiment 2 was to improve performance in the maintenance phase by manipulating the information listeners received in the learning phase and by giving the listeners explicit instructions regarding their categorizations (Condition 3). Both learning and maintenance stimuli were identical to those in Condition 1 of Experiment 1. Eighteen students of the University of Nijmegen participated in return for a small payment. The procedure of the maintenance phase was identical in all three conditions: listeners were asked to categorize the maintenance stimuli as they saw fit.

The learning phase differed according to condition. Condition 1 was identical to Condition 1 of Experiment 1: listeners received trial-by-trial right/wrong feedback on their responses. In Condition 2, listeners received right/wrong feedback on their responses and were provided with perceptual anchors (consisting of the means of the categories) 300 ms after each stimulus for the first 40 trials to aid in their categorization. In Condition 3, listeners received a written explanation as well as a visual depiction of the distributional properties of the stimuli (explaining the diagonal categorization rule and the importance of integrating the two stimulus dimensions to avoid errors), as well as the anchors and the feedback. Although explicit instruction has been been shown to hamper the acquisition of phonetic contrasts [11], it might be beneficial when combined with trial-by-trial feedback in the learning phase.

3.1. Results

The results presented in Table 3 show that trial-by-trial feedback nor additional perceptual anchors were helpful in facilitating multidimensional categorization in the maintenance phase. The Φ was 0 in both conditions and the number of listeners using both dimensions dropped to 1. Providing a verbal description while reminding the listeners of the importance of using both dimensions *and* providing trial-by-trial feedback *and* providing the anchors was helpful judging by the mean Φ and the number of listeners using both dimensions in the maintenance phase of Condition 3.

Table 3: *Mean Φ and A values and the number of listeners using both dimensions in Experiment 2.*

Cond.	N		Learning	N_{multi}	Maintenance	N_{multi}
1	6	Φ (σ)	0.24 (0.05)	6	0.06 (0.10)	1
		A (σ)	0.56 (0.22)		0.94 (0.63)	
2	6	Φ (σ)	0.25 (0.12)	4	-0.03 (0.07)	1
		A (σ)	0.35 (0.16)		1.03 (0.45)	
3	6	Φ (σ)	0.23 (0.11)	6	0.18 (0.11)	5
		A (σ)	1.36 (0.22)		0.70 (0.28)	

Statistical evaluation of the Φ's of the learning and maintenance phases confirmed these observations. In the learning phase of Conditions 1 and 2, Φ differed significantly from 0 and 0.5π (t_{min} = 5.2, p < 0.05). In the maintenance phase of these conditions, however, Φ only differed significantly from 0.5π (t_{min} = 11.2, p < 0.05) showing a return to unidimensional categorization. The learning phase of Condition 3 had similar results with Φ differing from both 0 (t [5] = 4,8, p < 0.05) and 0.5π (t [5] = -5,7, p < 0.05). In the maintenance phase, 3 Φ also differed significantly from both 0 (t [5] = 4.0, p < 0.05) and 0.5π (t [5] = -7.4, p < 0.05) indicating maintenance of the learned multidimensional categorization strategy.

The consistency measure A differed significantly from zero in all phases in all conditions (t_{min} = 3.7, p < 0.05). The ANOVA showed a statistical difference between the conditions in the learning phase (F [2,15] = 40.98, p < 0.05), but not in the maintenance phase (F [2,15] = 0.78, n.s.). Post hoc testing (Tukey HSD) confirmed that A is significantly higher in the learning phase of Condition 3 compared to Conditions 1 and 2, but not in the maintenance phase. The effect of the written explanation is thus present in the comparison of the Φ's, but not in the consistency measure A. In Conditions 1 and 2 listeners consistently used a multidimensional (learning phase) or a unidimensional (maintenance phase) categorization strategy, in Condition 3, they consistently used a multidimensional categorization strategy in all phases.

3.2. Discussion

Experiment 2 aimed at improving categorization performance in the maintenance phase by changing the learning conditions. In Condition 1, listeners received right/wrong feedback to help them learn to integrate the two relevant dimensions. Condition 2 added perceptual anchors to further help category learning. Finally, Condition 3 added a written instruction describing the category structures and stressed the importance of using both dimensions. Only with the all three manipulations could listeners maintain the categorization rule they successfully acquired in the learning phase.

4. Conclusions

Multidimensional category learning is possible after only a few hundred learning stimuli when trial-by-trial feedback and distributional properties are present. This replicates [12]. This learning does not persist once feedback and distributional information are no longer present. In [9] it was shown that learning is maintained if feedback is withheld but distributional information is of the stimulus set is kept, arguing for the importance of distributional information.

Making a unidimensional rule less effective during training by making the categories less distinct did not enhance the maintenance of multidimensional categorization. Experiment 2 shows that the assumptions listeners bring to the task (on their own initiative or through explicit instructions) can have a large influence on their categorization behavior.

Distributional learning is undoubtedly critical to category formation in many domains, including phonological development. At present, little is known about the specific characteristics and limitations of distributional learning, particularly for auditory categories. The present studies contribute to a growing body of evidence revealing limitations on the conditions under which listeners will maintain their learning once immediate feedback is withheld.

5. Acknowledgements

6. References

[1] L. Lisker, "Rapid vs. rabid: A catalog of acoustic features that may cue the distinction,", Haskins Laboratories, New Haven, CT, Status Report on Speech Research 54. 127-132, 1978.

[2] P. K. Kuhl, K. A Williams, F., Lacerda, K. N. Stevens, and B. Lindblom, B., "Linguistic experience alters phonetic perception in infants by 6 months of age," *Science*, vol. 255, pp. 606-608, 1992.

[3] L. Polka and J. F. Werker, "Developmental changes in perception of non-native vowel contrasts," *Journal of Experimental Psychology: Human Perception and Performance,* vol. 20, pp. 421-435, 1994.

[4] B. A. Younger, "The segregation of items into categories by 10-month old infants," *Child Development*, vol. 56, pp. 1574-1583.

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

[6] J. Maye and L. Gerken, "Learning phonemes: How far can the input take us?" in *Proceedings of the 25th Annual Boston University Conference on Language Development,* 2001, pp. 480-490.

[7] F. G. Ashby and W. T. Maddox, "Integrating information from separable psychological dimensions," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 16, pp. 598-612, 1990.

[8] F. G. Ashby, S. Queller, and P. Beretty, "On the dominance of unidimensioal rules in unsupervised categorization," *Perception & Psychophysics*, vol. 61, pp. 1178-1199, 1999.

[9] M. Goudbeek, "The acquisition of auditory categories," Ph.D. Dissertation, University of Nijmegen/MPI for Psycholinguistics, Nijmegen, The Netherlands, 2007.

[10] R. Smits, J. Sereno, and A. Jongman, "Categorization of sounds," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 32, pp. 733-754, 2006.

[11] M. Gulian, P. Escudero, and P. Boersma, "Supervision hampers distributional learning of vowel contrasts," submitted to *the 16th International Congress of the Phonetic Sciences*, Saarbrücken, Germany, 2007.

[12] M. Goudbeek and D. Swingley (2006), "Saliency effects in distributional learning," *Proceedings of the Eleventh Australasian International Conference on Speech Science and Technology,* Auckland, New Zealand, 2006.