Auditory categories with separable decision boundaries are learned faster with full feedback than with minimal feedback (L)

Han Gyol Yi and Bharath Chandrasekaran[a]

Department of Communication Sciences and Disorders, 2504A Whitis Avenue (A1100),
The University of Texas at Austin, Austin, Texas 78712, USA

During visual category learning, full feedback (e.g., "Wrong, that was a category 4."), relative to minimal feedback (e.g., "Wrong."), enhances performance when the relevant dimensions are separable. This pattern is reversed with inseparable dimensions. Here, the interaction between trial-by-trial feedback and separability of dimensions in the *auditory* domain is examined. Participants were trained to categorize auditory stimuli along separable or inseparable dimensions. One group received full feedback, while the other group received minimal feedback. In the separable-dimensions condition, the full-feedback group achieved higher accuracy than did the minimal-feedback group. In the inseparable-dimensions condition, performance was equivalent across the feedback groups. These results altogether suggest that trial-by-trial feedback affects auditory category learning performance differentially for separable and inseparable categories. © *2016 Acoustical Society of America.*
[http://dx.doi.org/10.1121/1.4961163]

[EAS]   Pages: 1332–1335

I. INTRODUCTION

Successful auditory categorization depends on the ability to form optimal boundaries across a perceptual space (Holt and Lotto, 2006). The boundaries can be either decisionally separable or inseparable. When the boundaries are decisionally separable [Fig. 1(a)], decisions along one dimension are independent from the values of another dimension (Garner, 1974). For example, loudness judgment of a pure tone stimulus is largely unaffected by the pitch of the tone, and vice versa (Grau and Nelson, 1988). In contrast, when the boundaries are decisionally inseparable [Fig. 1(a)], decisions along one dimension critically depend on the values of another dimension (Garner, 1974), as is the case for many speech sounds (Macmillan *et al.*, 1999; Hillenbrand *et al.*, 1995).

Currently, it is unclear how separability of optimal decision boundaries affects learning of auditory categories via trial-by-trial feedback. Research in vision suggests that an unfamiliar category structure is initially approximated using separable boundaries, which are often verbalizable (Ashby and Maddox, 2011). Learners are more likely to persist using separable boundaries when the feedback provides the correct category [e.g., "Wrong, that was a category 4."; "full feedback"; Fig. 1(b)], which facilitates explicit testing of the hypothesized decision boundaries, relative to when the feedback only indicates the accuracy of the response [e.g., "Wrong."; "minimal feedback"; Fig. 1(b); Maddox *et al.*, 2008]. Conversely, learners are more likely to switch to using inseparable boundaries when minimal feedback, rather than full feedback, is presented (Maddox *et al.*, 2008). Hence, categories based on separable boundaries are learned faster with full feedback, while those based on inseparable

boundaries are learned faster with minimal feedback (Maddox *et al.*, 2008). Since capturing the switch in strategy use is not a trivial task in a behavioral paradigm, manipulation of the feedback type (full vs minimal), combined with different category structures (separable vs inseparable), serves as a useful tool in characterizing the dynamics of auditory category learning.

There is some evidence that suggests that the initial bias for the separable boundaries is present during auditory category learning as well. When participants learn Mandarin Chinese lexical tone categories, where the optimal decision boundaries are posited to be inseparable (Chandrasekaran *et al.*, 2010; Maddox and Chandrasekaran, 2014; Chandrasekaran *et al.*, 2014a), adult learners initially attempt to identify the tone categories using separable boundaries but gradually transition to using inseparable boundaries (Maddox and Chandrasekaran, 2014; Chandrasekaran *et al.*, 2014a). Minimal feedback, relative to full feedback, leads to faster learning (Chandrasekaran *et al.*, 2014b). However, naturalistic categories, such as phonetic categories, can have multiple relevant perceptual dimensions (Lisker, 1986). For this reason, separability of the decision boundaries cannot be determined with certainty. Furthermore, in studies using naturalistic stimuli, the lack of a control category structure with known *separable* dimensions limits the extent to which feedback effects can be attributed to perceptual separability.

In the current study, we examined the extent to which separability of the optimal decision boundaries relates to feedback-driven auditory category learning. More specifically, we compared learning performance of separable and inseparable categories across full and minimal feedback to test the extent to which learners are initially biased towards using separable boundaries. Participants (N = 22) were trained on auditory categories based on two experimenter-

---

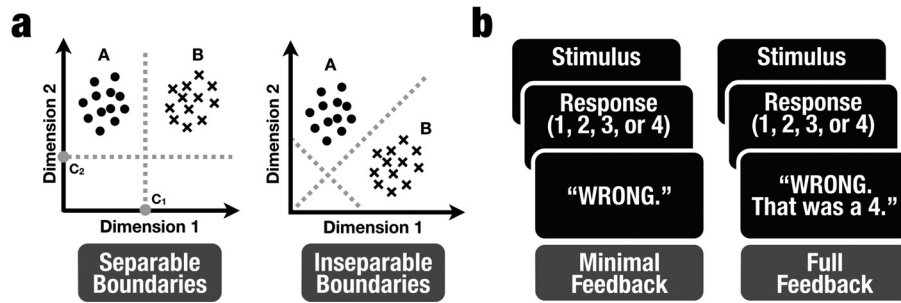[a]Electronic mail: bchandra@austin.utexas.edu

FIG. 1. (a) The optimal decision boundaries (dotted lines) separating a given set of categories (A and B) can be separable (left), meaning that they are reducible to a set of decisional criteria along the relevant dimensions (c_1 and c_2; i.e., a set of intercepts for each axis). In contrast, the optimal decision boundaries can also be inseparable (right), meaning that they are not reducible to a set of intercepts for each axis. (b) Categories can be learned in a supervised manner, using trial-by-trial corrective feedback. "Minimal" feedback indicates the accuracy of the response, but provides no other information. "Full" feedback, in addition to the accuracy of the response, also indicates what the correct response would have been.

constrained dimensions that are known to be important in audition: spectral modulation frequency and temporal modulation frequency (Singh and Theunissen, 2003). Across separate days, participants were presented with a category structure based on separable boundaries or inseparable boundaries. One group of participants (N = 11) exclusively received full feedback, while the other group (N = 11) received only minimal feedback. We hypothesized that results from visual category learning and speech category learning would be generalizable to the non-speech auditory category learning. As such, it was predicted that the full feedback group would *outperform* the minimal feedback group for the separable boundaries, but that the minimal feedback group would *outperform* the full feedback group for the inseparable boundaries.

II. METHODS

A. Participants

Young adult (18 to 35 yr of age; mean age = 22.8 yr, SD = 4.6 yr) native speakers of American English (N = 22; 13 females) with no self-reported hearing deficits (screened audiological thresholds <25 dB hearing level across 1, 2, and 4 kHz) were recruited from the University of Texas at Austin community. Participants received monetary compensation for their participation. The Institutional Review Board

at The University of Texas at Austin approved the recruitment, consent, testing, and payment procedures.

B. Stimuli

All sounds were generated by applying spectrotemporal modulation to a white noise stimulus (duration = 1 s; digital sampling rate = 44.1 kHz; low-pass filtered at 1.6 kHz). Spectral modulation frequency ranged from 0.1 to 2 cyc/oct. Temporal modulation frequency ranged from 4 to 12 Hz [Fig. 2(a)]. These modulation frequency ranges were based on the receptive field properties of the human auditory cortex (Schönwiesner and Zatorre, 2009). The amplitude of modulation depth was 30 dB. All sounds were RMS amplitude normalized to 82 dB. In order to create a category structure that can be optimally categorized using separable boundaries, 600 coordinates were first generated in an abstract two-dimensional space, with the minimum value of 0 and the maximum value of 1. Four bivariate normal distributions of 150 coordinates were centered on (0.33, 0.33), (0.33, 0.75), (0.75, 0.33), and (0.75, 0.75), with a standard deviation of 0.1. Values along each dimension were logarithmically mapped onto spectral and temporal modulation frequencies. Thereby, the optimal decision boundaries corresponded to a set of criteria defined by the spectral modulation frequency of 0.45 cyc/oct and the temporal modulation frequency of 6.93 Hz. The
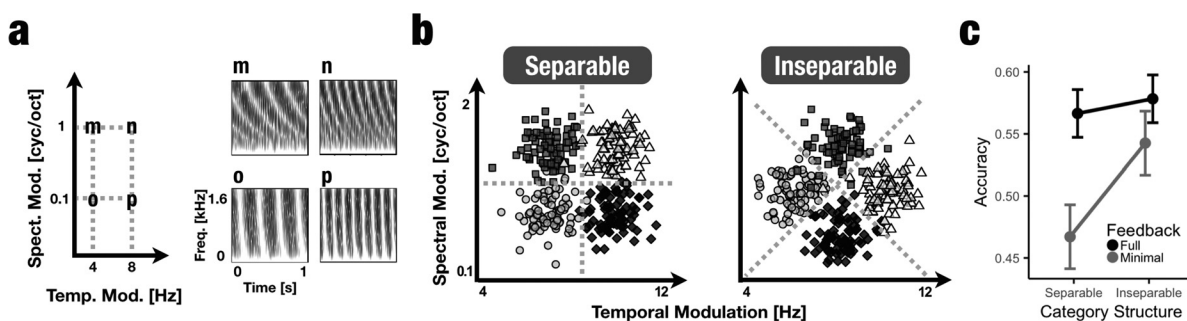

FIG. 2. (a) Spectrograms of example stimuli varying in spectral modulation frequency (0.1 and 1 cyc/oct) and temporal modulation frequency (4 and 8 Hz). (b) Two category structures were created based on the acoustic dimensions of temporal modulation frequency and spectral modulation frequency. The "separable" structure could be categorized using one decisional criterion on the spectral modulation dimension and another on the temporal modulation dimension. The "inseparable" structure was created by rotating the separable structure by 45° clockwise. (c) For the separable structure, categorization accuracy was higher for full feedback (black symbols and lines) relative to minimal feedback (gray symbols and lines). For the inseparable structure, categorization accuracy was equivalent across the feedback groups.

inseparable category structure was created by rotating the separable category structure by 45° clockwise [Fig. 2(b)].

C. Training procedures

Participants were placed in a sound attenuated booth, seated in front of a computer, and equipped with a pair of circumaural headphones. They were instructed to listen to each sound and categorize it into one of four categories using number keys 1, 2, 3, and 4 on a keyboard. Following a response, feedback was presented immediately. Each stimulus was presented once, resulting in 600 trials. The stimulus presentation order was randomized. In the full feedback group, feedback was provided in the form of, e.g., "Right, that was a category 4," or "Wrong, that was a category 4." In the minimal feedback group, feedback was presented as: "Right," or "Wrong."

D. Study design

Type of feedback was randomly varied in a between-subjects design, with a full feedback group (N = 11; 7 females; mean age = 22.1 yr, SD = 3.5 yr) or the minimal feedback group (N = 11; 6 females, mean age = 23.5 yr, SD = 5.6 yr). Type of category structure (separable versus inseparable) was varied in a within-subjects design, on separate days (gap range: 24–168 h), in a counter-balanced order.

E. Analysis

Separate analyses were performed for the separable and inseparable category structures. Logistic mixed effects modeling was performed with trial-by-trial accuracy as the dependent variable (Bates et al., 2013). Feedback type (minimal versus full) was set as the fixed effect, with the minimal feedback serving as the reference level for comparison. Therefore, two simple effects were estimated by the analysis. First, the feedback effect modeled the difference in log odds between the full feedback group and the minimal feedback group. A positive estimate for the feedback effect would suggest higher categorization accuracy for the full feedback group relative to the minimal feedback group, and vice versa. Second, the intercept modeled the log odds of a correct response for each participant receiving minimal feedback in each structure in a given trial. The null hypothesis would posit a zero estimate for the intercept, which corresponds to 50% categorization accuracy. Because the chance level in this task was 25%, we did not consider the intercept to be meaningful, but nevertheless it is reported for the sake of comprehensiveness, as all simple effects are estimated in reference to the intercept. The model was corrected for by-participant random intercepts and by-trial random slopes for each participant.

III. RESULTS

In the separable category structure, participants in the minimal feedback group correctly identified 47% (SD = 9%) of the stimuli, while the full feedback group correctly identified 57% [SD = 6%; Fig. 2(c), left]. Mixed effects analysis showed that the feedback effect was significant (b = 0.49, SE = 0.14, z = 3.43, p = 0.00061), suggesting that the

categorization accuracy was higher for the full feedback group than that for the minimal feedback group. The intercept was significant (b = −0.37, SE = 0.12, z = −3.10, p = 0.0020[1]). In the inseparable category structure, the participants in the minimal feedback group correctly identified 54% (SD = 9%) of the stimuli, while the full feedback group correctly identified 58% [SD = 6%; Fig. 2(c), right]. Mixed effects analysis showed that the feedback effect was not significant (b = 0.23, SE = 0.15, z = 1.49, p = 0.14), suggesting that the categorization accuracy was equivalent across the full and minimal feedback groups. The intercept was not significant (b = −0.05, SE = 0.12, z = −0.38, p = 0.71). In summary, the results for the separable category structure (full > minimal) followed our prediction (full > minimal). The results for the inseparable category structure (minimal = full) did not follow our prediction (minimal > full).

IV. DISCUSSION

The current study examined the extent to which the separability of the optimal decision boundaries interacts with manipulation of corrective feedback to affect non-speech auditory category learning. When the optimal decision boundaries were separable, full feedback, relative to minimal feedback, was associated with higher categorization accuracy. When the boundaries were inseparable, minimal and full feedback were associated with equivalent categorization accuracy.

Improved category learning performance for separable boundaries with full, relative to minimal, feedback has been previously demonstrated in the visual domain (Maddox et al., 2008). Separable boundaries can be easily verbalized as a set of decision criteria along the relevant dimensions. It has been suggested that the corrective information in full feedback ("Wrong, that was a category 4.") helps the learner evaluate alternative hypotheses that could have led to a correct response (Ashby and Maddox, 2011). In the auditory domain, the presence of feedback has been shown to enhance speech category learning in two investigations, but full and minimal feedback conditions could not be compared, as there were only two possible categories (Goudbeek et al., 2008; McCandliss et al., 2002). Full and minimal feedback conditions have been compared in speech category learning with four categories, but generalizability was limited, as the relevant dimensions had not been constrained by the experimenters (Chandrasekaran et al., 2014b). The current study is the first to suggest that non-speech auditory categories with separable optimal decision boundaries can be learned faster with full, relative to minimal, feedback.

Unexpectedly, learning performance did not differ in terms of full or minimal feedback for inseparable boundaries. This finding was inconsistent with the prediction that full, relative to minimal, feedback would be associated with lower learning performance. In the visual domain, it has been suggested that full feedback discourages learners from switching to using inseparable boundaries, rather than separable boundaries, to approximate the category structure (Maddox et al., 2008). The disruptive effect of full feedback has been replicated in speech category learning, where the

optimal decision boundaries are posited to be inseparable (Chandrasekaran *et al.*, 2014a; Chandrasekaran *et al.*, 2014b; Maddox and Chandrasekaran, 2014). One possible explanation for the equivalent performance between full and minimal feedback conditions in the auditory domain is that the learners were less likely to be initially biased towards using separable boundaries, when learning less easily verbalizable non-speech auditory categories than with speech auditory categories.

This article presents evidence that separability of decision boundaries interacts with the content of corrective feedback during non-speech auditory category learning. What remains unclear is the extent to which equivalent performance across feedback groups during inseparable category learning in audition is attributable to the lack of initial bias towards separable boundaries. This issue warrants further exploration using different sets of auditory stimuli with varying degrees of learning difficulty.

ACKNOWLEDGMENTS

[1]A separate computational modeling analysis was conducted to test the extent to which participants relied on a single dimension to categorize the separable stimuli. Sixty five percent of the model fits were multidimensional, suggesting that, despite the low accuracies, the participants were utilizing both dimensions.

Ashby, F. G., and Maddox, W. T. (**2011**). "Human category learning 2.0," Ann. N. Y. Acad. Sci. **1224**, 147–161.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (**2013**). "lme4: Linear mixed-effects models using Eigen and S4," R package version 1.

Chandrasekaran, B., Koslov, S. R., and Maddox, W. T. (**2014a**). "Toward a dual-learning systems model of speech category learning," Front. Psychol. **5**, 825.

Chandrasekaran, B., Sampath, P. D., and Wong, P. C. (**2010**)."Individual variability in cue-weighting and lexical tone learning," J. Acoust. Soc. Am. **128**, 456–465.

Chandrasekaran, B., Yi, H., and Maddox, W. T. (**2014b**). "Dual-learning systems during speech category learning," Psychon. Bull. Rev. **21**, 488–495.

Garner, W. R. (**1974**). "The stimulus in information processing," in *Sensation and Measurement* (Springer, Netherlands).

Goudbeek, M., Cutler, A., and Smits, R. (**2008**). "Supervised and unsupervised learning of multidimensionally varying non-native speech categories," Speech Commun. **50**, 109–125.

Grau, J. W., and Nelson, D. K. (**1988**). "The distinction between integral and separable dimensions: Evidence for the integrality of pitch and loudness," J. Exp. Psychol.-Gen. **117**, 347–370.

Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (**1995**). "Acoustic characteristics of American English vowels," J. Acoust. Soc. Am. **97**, 3099–3111.

Holt, L. L., and Lotto, A. J. (**2006**). "Cue weighting in auditory categorization: Implications for first and second language acquisition," J. Acoust. Soc. Am. **119**, 3059–3071.

Lisker, L. (**1986**). "'Voicing' in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees," Lang. Speech. **29**, 3–11.

Macmillan, N. A., Kingston, J., Thorburn, R., Dickey, L. W., and Bartels, C. (**1999**). "Integrality of nasalization and F1. II. Basic sensitivity and phonetic labeling measure distinct sensory and decision–rule interactions," J. Acoust. Soc. Am. **106**, 2913–2932.

Maddox, W., and Chandrasekaran, B. (**2014**). "Tests of a dual-system model of speech category learning," Biling.-Lang. Cogn. **17**, 709–728.

Maddox, W. T., Love, B. C., Glass, B. D., and Filoteo, J. V. (**2008**). "When more is less: Feedback effects in perceptual category learning," Cognition **108**, 578–589.

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., and McClelland, J. L. (**2002**). "Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception," Cogn. Affect. Behav. Neurosci. **2**, 89–108.

Schönwiesner, M., and Zatorre, R. J. (**2009**). "Spectro-temporal modulation transfer function of single voxels in the human auditory cortex measured with high-resolution fMRI," Proc. Nat. Acad. Sci. **106**, 14611–14616.

Singh, N. C., and Theunissen, F. E. (**2003**). "Modulation spectra of natural sounds and ethological theories of auditory processing," J. Acoust. Soc. Am. **114**, 3394–3411.