PS-158

Central Unmasking in a Melody Recognition Task

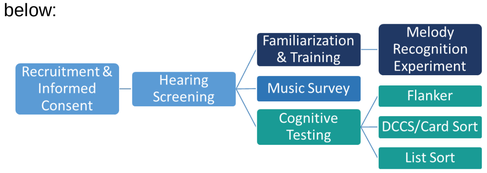
Brittany N. Jaekel, Corey Stoelb, Sara M. Misurelli, and Ruth Y. Litovsky
University of Wisconsin-Madison, USA
e-mail: jaekel@wisc.edu

37th Midwinter Meeting of the Assoc. Res. Otol.
San Diego, CA
February 22-26, 2014

INTRODUCTION

The extent to which we can extract melody and segregate source information in the presence of tonal maskers is not well known. Determining the listening environments in which melody recognition is improved or degraded in normal-hearing (NH) listeners, and how contralateral unmasking and divided attention affect melody recognition, will provide parameters for comparison with cochlear implant (CI) users. CI users typically receive degraded spectral information via their implants, which may impact their performance for certain conditions of the melody recognition task.

Kidd et al. (1998) reported increases in recognition accuracy of simple tonal sequences in the presence of informational maskers, with spatial separation of sources. Here, we used familiar melody sequences with tonal maskers in the hopes that results would better generalize to actual music listening experiences, and to facilitate training. Further, we employed cognitive and music experience measures to determine what factors may help or hinder recognition abilities. The sequence of the experimental tasks is shown in the figure below:



AUDITORY MEASURES

Participants

- 9 Normal Hearing (NH) adults (female = 6).
- Completed a **music experience survey**:

Self-Reported "Non-Musical", means

N	*Age*	*# Instruments*	*# Years Active*	*# Years since Active*
3	24	2	9.3	8

Self-Reported "Musical", means

N	*Age*	*# Instruments*	*# Years Active*	*# Years since Active*
6	22.8	3.2	22	1.2

Stimuli

- Stimuli were delivered via circum-aural headphones to the participant's right ear.
- Targets were the first seven notes of eight familiar melodies, presented as pure tones. Inclusion criteria for melody selection were 1) a beginning phrase of notes of equal (or nearly equal) length, and 2) ability to be transcribed to fit completely within the A_3-A_4 octave.
- Maskers were pure tone "distractor notes" pulled randomly from the same octave as the target melody.

Conditions:

Signal-to-Noise Ratio (SNR)	-15, -10, -5, 0, 5, 10 dB
Masker Temporal Configuration	• **Interleaving** with melody • **Overlapping** with melody • **Both** interleaving & overlapping with melody
Masker Location	• **Same** ear as melody (right) • **Opposite** ear from melody (left)

AUDITORY MEASURES, cont.

- Frequencies from base target melodies and maskers were multiplied by a random factor between 0.5 and 2. This was repeated for each testing run to reduce possibility of memorized pitch cues for particular melodies.

- Notes for base target melodies were equally spaced, and played for an equal length of time. This reduced possibility of rhythmic cues for particular melodies.

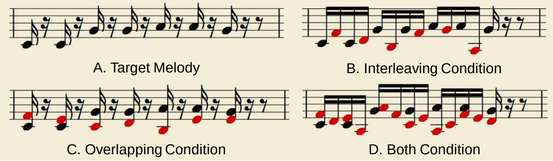


Figure 1. An example of each condition, using musical notation and the target melody "Twinkle, Twinkle, Little Star" in treble clef. Red notes indicate maskers.

Target Melodies	Base notes	Frequency Range
1. Old MacDonald Had a Farm	F_4-F_4-F_4-C_4-D_4-D_4-C_4	261.63 – 349.23 (Hz)
2. Twinkle, Twinkle, Little Star / ABC	C_4-C_4-G_4-G_4-A_4-A_4-G_4	261.63 – 440.00 (Hz)
3. Frère Jacques	C_4-D_4-E_4-C_4-C_4-D_4-E_4	261.63 – 329.63 (Hz)
4. London Bridge is Falling Down	G_4-A_4-G_4-F_4-E_4-F_4-G_4	329.63 – 440.00 (Hz)
5. My Country 'tis of Thee	C_4-C_4-D_4-B_3-C_4-D_4-E_4	246.94 – 329.63 (Hz)
6. Ode to Joy	E_4-E_4-F_4-G_4-G_4-F_4-E_4	329.63 – 392.00 (Hz)
7. Mary had a Little Lamb	E_4-D_4-C_4-D_4-E_4-E_4-E_4	261.63 – 329.63 (Hz)
8. Yankee Doodle	C_4-C_4-D_4-E_4-C_4-E_4-D_4	261.63 – 329.63 (Hz)

Table 1. List of the target melodies, along with their note names and frequencies (Hz). Melodies were originally composed using the octave A_3-A_4, as indicated by blue keys. Masker notes were also drawn from this octave.

Procedure

- 8-AFC, with melodies randomized with replacement. Participants pressed a button to initiate each run, and then chose the name of the melody they heard.

- Participants completed 114 trials. ([2 ear positions x 3 masker configurations x 6 SNRs x 3 trials each] + 6 trials of target-in-silence). Conditions were randomized.

Training

- Participants were trained on the task and had to reach 80% accuracy or higher on all masker configurations and locations presented at 10 SNR before beginning the experiment.

- Feedback and a repeat button were present during training, but removed during actual testing. During training, a correct melody choice turned the button green. An incorrect choice turned the correct melody's button green, and then played the correct melody in silence to the right ear.

RESULTS: MELODY RECOGNITION

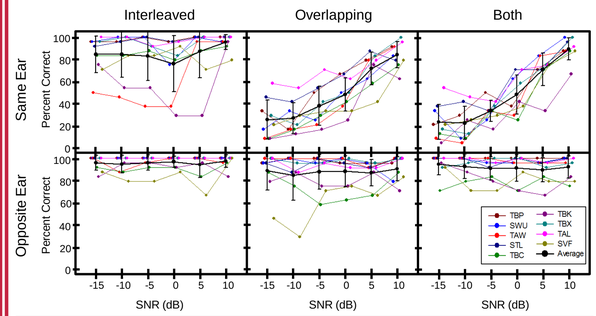


Figure 2. Individual participants' %correct melody recognition plotted as a function of SNR; averages (±1s.d.) are plotted in black.

- Largest individual differences were observed in the Same-Ear/interleaved condition. The largest effects as a function of SNR were seen in the Same-Ear conditions containing overlap in time. The effect of adding the interleaved to the temporal overlap did not produce additional masking (compare 'overlapping' with 'both').

RESULTS: MUSICAL BACKGROUND

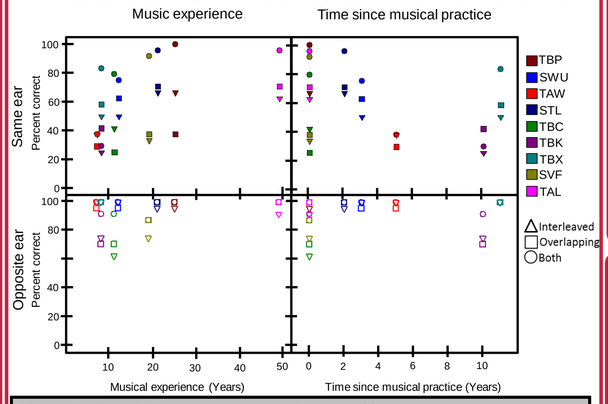


Figure 3. Correlations between melody recognition accuracy and either total years of active music training or recency of music activity are shown. Data consist of Same- or Opposite-Ear conditions at 0 dB SNR for each participant.

N=9		Same Ear			Opposite Ear		
		Interleaving	Overlapping	Both	Interleaving	Overlapping	Both
Recency	Pearson corr.	-.610	-.423	.066	.098	.207	-.123
	Sig. (2-tailed)	.081	.257	.866	.802	.594	.753
Years Active	Pearson corr.	.601	.610	.494	.229	.110	.411
	Sig. (2-tailed)	.087	.081	.176	.553	.777	.272

- In Same-Ear conditions, low recognition accuracy positively correlated with years of experience, and negatively correlated with recency (years since actively training/playing), for "interleaved" (triangles) and "both" (circles) conditions, but these correlations were not significant (see table).
- Accuracy at 0 dB SNR was chosen as representative of the participants' "strategy" for listening. A participant relying on a loudness difference cue would have a dip in performance at this SNR; a participant anticipating an "every-other" pattern of notes would show no change in performance at this SNR. Awareness of note order may be associated with musical training.

RESULTS: COGNITIVE

Working Memory

List Sort: Participants were presented with a series of items (food and/or animals in random size order) and instructed to verbally repeat the items in size order from smallest to largest. When both animals and food were shown within the same trial, the participant was instructed to repeat the food items from smallest to largest, then the animals.

Attention

Flanker: (Inhibitory control): Participants were shown a series of arrows and indicated the direction of the middle arrow, while simultaneously inhibiting the other arrows.

Dimensional Change Card Sort (DCCS): (Cognitive flexibility): Target pictures that varied along two dimensions (color, shape) were presented. Participants matched target pictures to test pictures for either color or shape, depending on if the Prompt was "shape" or "color".

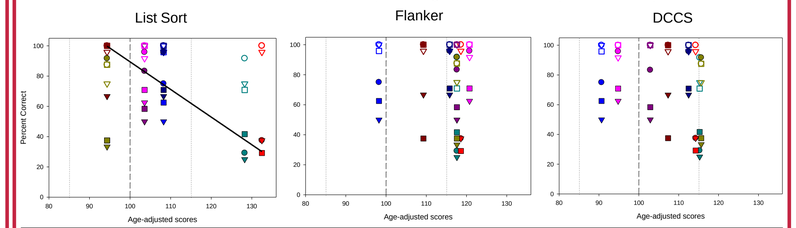
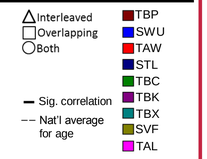


Figure 4. Correlations between cognitive scores and melody recognition accuracy. Data consist of Same- or Opposite-Ear conditions at 0 dB SNR for each participant.

N=8		Same Ear			Opposite Ear		
		Interleaving	Overlapping	Both	Interleaving	Overlapping	Both
ListSort	Pearson corr.	-.928**	-.541	-.285	.052	.004	-.459
	Sig. (2-tailed)	.001	.166	.494	.902	.993	.252
Flanker	Pearson corr.	-.134	-.206	-.123	-.259	-.355	-.156
	Sig. (2-tailed)	.751	.624	.771	.535	.388	.712
DCCS	Pearson corr.	-.357	-.463	-.633	-.548	-.507	-.473
	Sig. (2-tailed)	.386	.247	.092	.160	.199	.236



- The negative correlation of scores on the List Sort task and the Same-Ear interleaved condition was significant ($r = -.928$, $p < .01$).
- Mean group scores for our sample were slightly higher than normed average. Performance on the List Sort was generally variable, indicating possible differences in working memory across participants. Larger sample sizes may be needed to understand this more deeply.

CONCLUSIONS

- Contralateral separation of masker and target sources improves ability to extract melody in complex listening situations, if the participant is able to attend to the target ear. Conversely, overlapping (and, for some participants, interleaving) a masker sequence with the target melody degrades performance, perhaps due to listener uncertainty about which stream to attend to.
- Years of experience and recency of musical training are correlated positively & negatively, respectively, with same-ear performance in "interleaved" and "overlapping" conditions, though these correlations only approach significance. These correlations largely disappear in the contralateral masking conditions. Thus, musical training may be particularly beneficial for certain listening environments.
- **Future Research**: Comparisons with recognition abilities of users of CIs. Do they experience the same informational unmasking effects? Could training on certain musical aspects like pattern anticipation enhance performance?

REFERENCES

Kidd, Jr., G., Mason, C. R., & Arbogast, T. L. (2002). Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns. *J. Acoust. Soc. Am., 111*(3), pp. 1367-1376.
Kidd, Jr., G., Mason, C. R., Rohtla, T. L., & Deliwala, P. S. (1998). Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. *J. Acoust. Soc. Am.,* 104(1), pp. 422-431.
Levitin, D. J. (2012). What does it mean to be musical? *Neuron, 73*, pp. 633-637.
National Institutes of Health & Northwestern University. (2012). *NIH Toolbox for the Assessment of Neurological and Behavioral Function*. Retrieved Jan 2014, from www.nihtoolbox.org.
Patel, A. D. (2010). *Music, Language, and the Brain*. USA: Oxford University Press.

ACKNOWLEDGEMENTS

Work supported by NIH-NIDCD (R01-DC003083; Litovsky PI) and NIH-NICHD P30HD03352 (Waisman Center Core Grant).