

# Perception of Prosodic Cues by Japanese EFL Learners

Kazuhito Yamato

Kobe University

Shinobu Mizuguchi

Kobe University

## Reference Data:

Yamato, K., & Mizuguchi, S. (2014). Perception of prosodic cues by Japanese EFL learners. In N. Sonda & A. Krause (Eds.), *JALT2013 Conference Proceedings*. Tokyo: JALT.

This paper is a report on an investigation of how Japanese EFL learners perceive prosodic cues in spoken English, based on a series of studies using *Rapid Prosody Transcription* tasks (Cole, Mo, & Baek, 2010; Mo & Cole, 2010; Mo, Cole, & Lee, 2008). Thirty-eight Japanese undergraduate students participated in the experiment. Based on their TOEFL scores, a higher and a lower proficiency group were extracted. The strong inter-listener agreement in boundary perception for both groups suggests pauses were the most important cue for boundary detection, which is analogous to the perception by native speakers in Mo et al. (2008). However, both groups showed only weak inter-listener agreement in prominence perception. Prominence markings showed weak inter-listener agreement, and proficiency differences for boundary and prominence perception were not observed.

本研究は、Mo, Cole, & Lee (2008)らによって開発された“Rapid Prosody Transcription” tasksという、音声を聞きながらスクリプトにチャックごとに境界線を引く、卓立 (prominence) と思われる語句に下線を引く課題を、日本人大学生38人に実施し、どのようにプロソディ要素を知覚するのかを調査したものである。また、TOEFLスコアを基に上位群・下位群の2群を設定し、習熟度による違いを検討した。境界の知覚については、Mo et al. (2008)の英語母語話者と似た傾向があり、ポーズを主な手がかりとしていることが明らかとなった。卓立の知覚については、被調査者間で一致度が低く、幅広い語句を卓立として選択している傾向が見られた。習熟度による違いは特に見られなかった。

**A**MONG THE various elements of pronunciation, the appropriate production of prosody may be the most critical aspect for both speakers and listeners in interaction. Although errors in segmentals can lead to serious misunderstandings in communication, the contribution of prosody to intelligibility and successful communication deserves equal or even greater attention (Trofimovich & Baker, 2006; Gilbert, 2008). To interpret an intended message, the listener must use much of the prosodic information in an utterance. Listeners detect the prosodic boundaries encoded in speech to break utterances into chunks of words and identify prosodic prominences that serve to highlight a word or phrase guiding the listener to the most important information. Thus, prosodic information plays an important role in the listener's comprehension of utterances.

There has been growing interest in teaching and learning the prosodic features of English under the influence of Communicative Language Teaching (CLT; Celce-Murcia, Brinton, & Goodwin, 2010; Pennington & Richards, 1986). However, the use of prosodic features by Japanese learners of EFL deserves further investigation (see Ueyama, 2000; Yamato, 2004). There-



fore, in order to improve EFL teaching strategies and to provide a better learning environment for the acquisition of prosodic features, it is crucial to start with an understanding of how Japanese EFL learners perceive prosodic features of English.

Mo, Cole, and Lee (2008) researched the perception of prosodic features by average native speakers of English. They examined how untrained listeners (i.e., listeners with no phonetic transcription training) perceive prosodic boundaries and prominence in spontaneous speech taken from the Buckeye corpus (Pitt et al., 2007) by using real-time transcription tasks called *Rapid Prosody Transcription (RPT) tasks*. In these tasks, participants listen to recorded speech and mark the location of boundaries and prominent words on a printed transcript. The Mo et al. (2008) study showed that even these listeners responded in a consistent manner and can provide invaluable information on how they perceive prosodic boundaries and prominence.

The Mo et al. (2008) study raised important issues when the approach is adapted for Japanese EFL learners, which lead to the research questions for the present study:

1. Do Japanese EFL learners consistently perceive two types of prosodic cues (i.e., boundaries and prominence) in natural speech?
2. Does their proficiency affect their perception?

To answer these questions, we investigated how Japanese EFL learners perceive prosodic cues in spontaneous speech.

## Methodology and Analysis

### Materials

The original 20-30 second excerpts of conversational speech taken from the Buckeye corpus used by Mo et al. (2008) were too difficult for the participants in this study. To resolve this problem, we created two shorter sets of stimuli (5-10 seconds

long) from the Buckeye corpus that were relatively easy to comprehend.

Sound files created for this study (33 in total) were arranged into two blocks: a prominence transcription block with 11 audio stimuli and a boundary block with 12 stimuli. The participants received a printed transcript of each excerpt, with words separated by a space with no punctuation or capitalization (see Appendix).

### Participants and Transcription Procedure

The survey participants were Japanese EFL learners ( $N = 38$ ) recruited from undergraduate English classes at Kobe University. Their average TOEFL PBT score was 483.5 ( $SD = 25.86$ ). Based on these scores, two groups were extracted, a higher level group (H-JEFLs,  $n = 10$ ,  $x > \text{mean} + 0.5 SD$ ) and a lower level group (L-JEFLs,  $n = 11$ ,  $x < \text{mean} - 0.5 SD$ ).

The transcription tasks consisted of a boundary task and a prominence perception task. The survey procedure followed the experiment design outlined in Mo et al. (2008). After listening to a brief introduction and granting their informed consent, the participants marked the printed transcripts for prominence and boundaries of speech chunks as they listened to the recorded stimulus.

Prominent words, explained in the task sheet as words that have a highlighting function for the listener and stand out from the surrounding words, were to be marked by underlining. The boundaries of speech chunks, explained in the task sheet as a grouping of words “that helps the listener interpret the utterance” (Mo et al., 2008, p. 736), were to be marked by vertical lines. The participants were free to revise their responses as each speech excerpt played a second time.

## Analyses

The transcribed data analysis consisted of an inter-listener agreement in boundary and prominence perception task. The second analysis focused on the probabilistic scores for boundary and prominence perception by listeners overall, that is, both the higher and lower level groups. The number of possible responses normalized the raw numbers for boundary and prominence marks. The resultant measurement is the probability that a given place will receive a boundary or prominence mark, called *b-score* and *p-score*, respectively.

## Results

### Inter-Listener Agreement

The levels of agreement on boundary and prominence markings were evaluated using Fleiss' kappa coefficient (see Mo et al., 2008), which assesses the reliability of agreement between all pairs of listeners.

**Table 1. Inter-Listener Agreement of Japanese EFL Learners on Boundary and Prominence Marking (Kappa Coefficient)**

Marking	Overall	H-JEFLs	L-JEFLs
Boundary	.676	.704	.670
Prominence	.272	.260	.277

In Table 1, the results indicate higher than chance agreements among the listeners for boundary perception. For prominence perception, however, the inter-listener agreements were not so high. There was not much difference between H-JEFLs and L-JEFLs in terms of either their boundary or prominence perception.

## Boundary Perception

As shown in Table 1, both H-JEFLs and L-JEFLs showed a high level of agreement in boundary perception. Furthermore, it is clear from the visual representations in Figures 1 and 2 that the b-score distributions are similar for the two groups, which indicates both groups likely detected the same cues in their boundary perception.

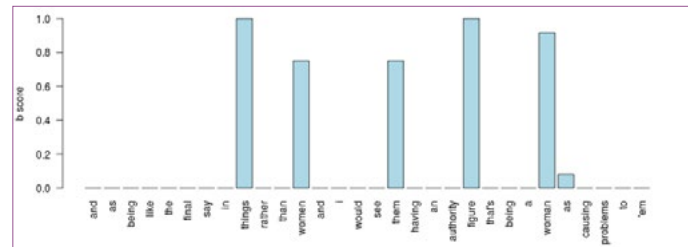


Figure 1. B-scores of L-JEFLs on audio stimulus #2.

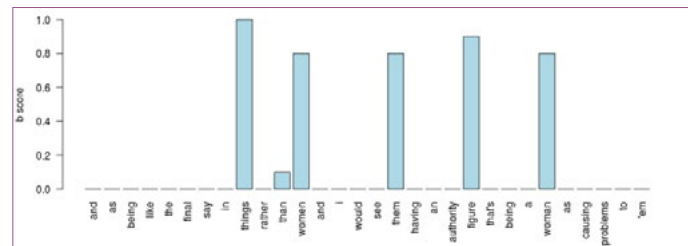


Figure 2. B-scores of H-JEFLs on audio stimulus #2.

## Prominence Perception

As was shown in Table 1, both groups showed a low level of agreement in prominence perception. As can be seen in the

visual representations in Figures 3 and 4, the distributions of *p*-score are not consistent for H- and L-JEFLs, which implies that Japanese EFL learners overall do not rely on prominence as a prosodic cue when listening to English.

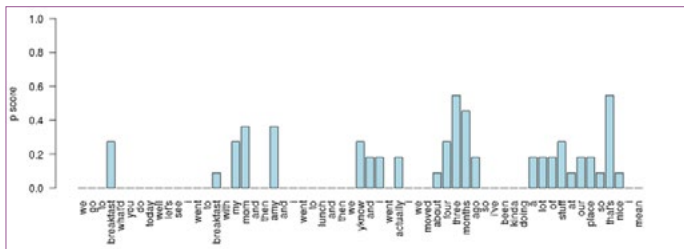


Figure 3. *P*-scores of L-JEFLs on audio stimulus #5.

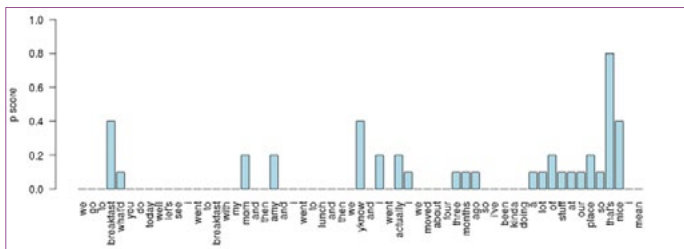


Figure 4. *P*-scores of H-JEFLs on audio stimulus #5.

## Discussion

### Inter-Listener Agreement

As Table 1 shows, the agreement score obtained for boundary perception revealed that both H- and L-JEFL learners are consistent in their labeling of boundaries. Furthermore, boundary perception is more consistent than prominence perception for both groups. However, the low agreement score in prominence

perception as opposed to boundary perception does not seem to be a phenomenon unique to Japanese EFL learners, as it has also been reported in several RPT studies (Cole, Mo, & Baek, 2010; Mo et al., 2008; Smith, 2011; Smith & Edmunds, 2013).

**Table 2. Inter-Listener Agreement of Native Speakers of English on Boundary and Prominence Marking (Kappa Coefficient)**

Marking	Group 1	Group 2	Group 3	Group 4
Boundary	.612	.544	.621	.575
Prominence	.373	.421	.394	.407

*Note.* Adapted from Mo, Cole, and Lee (2008).

Table 2 shows the inter-listener agreement on boundary and prominence marking by 74 native speakers of English taken from Mo et al. (2008). These data are for reference only because the materials used in the Mo et al. study are different from those of the present study. (Note, however, that the results of native speakers of English look similar to those of the Japanese EFL learners in Table 1). Both native speakers and Japanese EFL learners have a high level of agreement in boundary perception. The prominence score of Japanese EFL learners, however, seems relatively lower than that of native speakers of English.

### Boundary Perception

As the results in Table 1 show, both H- and L-JEFL learners are consistent in their labeling of boundaries, which is analogous to the results of native speakers of English in Mo et al. (2008).

Mo and Cole's (2010) research results showed that pause is the primary correlate for boundary perception by native speakers of

English, although other cues aid listeners to identify boundaries as well. Is pause also a primary cue for boundary perception by Japanese EFL learners? To address this question, another step was taken in which the normalized pause durations of audio stimuli were superimposed on b-score data. As Figure 5 clearly shows, b-scores correlated with pauses.

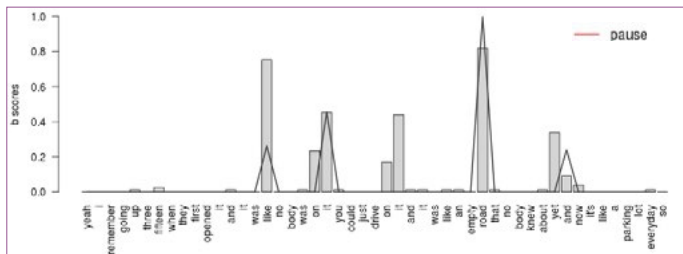


Figure 5. B-scores and normalized pause duration.

However, boundary responses did not completely coincide with pauses. Pinter, Mizuguchi, and Yamato (2014) reported that boundary responses in their study might have been triggered by prolonged vowels (e.g., as a sign of hesitation), discourse fillers such as *uhm* or *eh*, and slower tempos.

In addition, syntactic information can be another cue affecting boundary perception. Cole et al. (2010) argued that syntactic context plays an important role in boundary perception for native speakers of English. Pinter et al. (2014) reported that Japanese EFL learners also perceive boundaries by relying not only on pauses, but on the more frequent minor phrases as well, such as clause endings and noun phrases (see Kawahara & Shinya, 2008). Further research is called for to specify acoustic cues for boundary perception.

## Prominence Perception

Listeners recognize prominence phonetically as a combination of pitch, duration, and intensity (e.g., Beckman, 1986; Bolinger, 1986; Cruttenden, 1997; Pierrehumbert, 1980). The evidence is not yet conclusive on whether or not pitch functions as a (or perhaps the only) primary cue for prominence in English (Beckman, 1986; Pierrehumbert, 1980). Mo (2008), for example, found that native speakers of English utilized acoustic cues such as duration and intensity as well as pitch for prominence detection. Watanabe (1988), on the other hand, reported that Japanese EFL learners tended to rely on pitch as a cue for prominence rather than vowel duration and intensity and suggested the presence of an L1 influence.

In the present study, a lower agreement score on prominence perception exists for both proficiency groups of Japanese EFL learners. This result suggests Japanese EFL learners' inconsistency in perceiving acoustic cues for prominence. Following Watanabe's (1988) argument, we may speculate that Japanese EFL learners in this study also relied on pitch as a primary cue for prominence, rather than on vowel duration and intensity. Further research is required to explore the effects of the combination of acoustic cues on prominence perception and the effects of L1 interference.

## Implications for Teaching Listening and Pronunciation

We can draw some implications for teaching listening and pronunciation to Japanese EFL learners based on the results of this study. First, the Japanese EFL learners in this study showed high levels of agreement in boundary perception regardless of their proficiency levels. However, they utilized mainly silent pauses as a prosodic cue, unlike the native speakers of English in Mo et al. (2008), who used combinations of pauses and syntactic

cues. This is not a surprising tendency, as the detection of silent pauses is a universal feature, but the issue deserves attention in EFL instruction. When teaching listening and pronunciation, instructors should explain the relationship between pause and boundary perception, as bearing this relationship in mind can be a useful strategy for learners.

Second, low levels of agreement in prominence perception were obtained for both proficiency groups of Japanese EFL learners. The reasons behind this phenomenon might be L1 interference or lack of knowledge about prominence placement, which clearly requires further research.

These findings suggest that EFL instructors should design listening and pronunciation activities that incorporate the basic principle of prominence placement as a prosodic cue, which is “one nucleus in a phrase boundary” (Saito & Ueda, 2011; see also Celce-Murcia et al., 2010; Hahn, 2004; Nanjo, 2010).

### Concluding Remarks

In this study we investigated if Japanese EFL learners consistently perceived English prosodic cues in natural conversational speech in an RPT task, and if their proficiency affected their perception.

Prominence perception by Japanese EFL learners, regardless of their proficiency level, had low agreement, however agreement in boundary detection was relatively high, but no salient difference between the group with higher proficiency and the group with lower proficiency. In the case of boundary perception, the Japanese EFL learners seemed to detect pauses as the most significant prosodic cues. However, they seemed unable to apply their prominence perception strategies to English speech or not to know about prominence. These results have some implications for listening and pronunciation instruction, such as the need to teach bridging syntactic structures with prosodic features and understanding the basic principle of nucleus placement.

As a direction for further research, a more detailed analysis of boundary and prominence perception is called for to clarify the relationship between syntactic structures and prosodic features. This will allow for a more detailed description of Japanese EFL learners’ perception behavior and eventually yield valuable insights into EFL teaching strategies. Additionally, it is necessary to investigate the perception behavior of a wider range of English proficiency levels, as the participants in this study had only a narrow range of proficiency. Another direction for further research would be to explore how Japanese EFL learners utilize prosodic cues in their production of natural conversational speech in English and how native or nonnative speakers perceive their speech.

### Acknowledgments

We would like to express our sincere thanks to Professor Jennifer Cole for kindly letting us access the sound stimuli, Professor Gabor Pinter for assisting with the statistical analysis, and all the participants in this survey. The JSPS KAKENHI Grant #24520542 supported this research.

### Bio Data

**Kazuhito Yamato** is an associate professor at the School of Languages and Communication, Kobe University. His research interests include L2 pronunciation instruction for Japanese EFL learners, acquiring/learning English prosody, and interlanguage pragmatics. <yamato@port.kobe-u.ac.jp>

**Mizuguchi Shinobu** is a professor at the Graduate School of Intercultural Studies, Kobe University. Her research interests include formal semantics and L2 prosody learning. <mizuguti@kobe-u.ac.jp>

## References

- Beckman, M. E. (1986). *Stress and non-stress accent*. Dordrecht, Germany: Foris Publications.
- Bolinger, D. (1986). *Intonation and its parts*. Stanford, CA: Stanford University Press.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (with Griner, B.). (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). New York: Cambridge University Press.
- Cole, J., Mo, Y., & Baek, S. (2010). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes*, 25, 1141-1177.
- Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge: Cambridge University Press.
- Gilbert, J. B. (2008). *Teaching pronunciation: Using the prosody pyramid*. Cambridge: Cambridge University Press.
- Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38, 201-223.
- Kawahara, S., & Shinya, T. (2008). The intonation of gapping and coordination in Japanese: Evidence for intonational phrase and utterance. *Phonetica*, 65, 62-105.
- Mo, Y. (2008). Acoustic cues of prosodic prominence for naïve listeners of American English. *Proceedings of the 34th Meeting of the Berkeley Linguistic Society*. Retrieved from [http://prosody.beckman.illinois.edu/pubs/42\\_Mo\\_2008\\_bls34.pdf](http://prosody.beckman.illinois.edu/pubs/42_Mo_2008_bls34.pdf)
- Mo, Y., & Cole, J. (2010). Perception of prosodic boundaries in spontaneous speech with and without silent pauses. *The Journal of the Acoustical Society of America*, 127, 1956.
- Mo, Y., Cole, J., & Lee, E. (2008). Naïve listeners' prominence and boundary perception. In P. A. Barbosa, S. Madureira, & C. Reis (Eds.), *Proceedings of the Fourth International Conference on Speech Prosody* (pp. 735-738). Campinas, Brazil: ISCA. Retrieved from <http://www.isca-speech.org/archive/sp2008>
- Nanjo, K. (2010). Onseigaku oninron to hatsuonshidou [English phonetics and phonology and teaching pronunciation]. In N. Okada, K. Minamide, & A. Umesaki (Eds.), *A series of studies and English education. Volume 8, English studies and English education—Applying results of English studies to English education* (pp. 3-21). Tokyo: Taishukan.
- Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly*, 20, 207-225.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation* (Doctoral dissertation). Massachusetts Institute of Technology. Retrieved from <http://faculty.wcas.northwestern.edu/~jbp/publications.html>
- Pinter, G., Mizuguchi, S., & Yamato, K. (2014). Boundary and prominence perception by Japanese learners of English: A preliminary study. *Phonological Studies*, 17, 59-66.
- Pitt, M. A., Dille, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. (2007). *Buckeye corpus of conversational speech* (2nd release). Columbus: Ohio State University (Distributor). Available from <http://www.buckeyecorpus.osu.edu>
- Saito, H., & Ueda, I. (2011). Eigogakushusha ni yoru intoneshonkaku no gohaichi [Misplacement of nuclear stress by Japanese learners of English]. *Journal of the Phonetic Society of Japan*, 15, 87-95.
- Smith, C. (2011). Perception of prominence and boundaries by naïve French listeners. In W.-S. Lee & E. Zee, *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1874-1877). Hong Kong: ICPhS.
- Smith, C., & Edmunds, P. (2013). Native listeners' perception of prosody in L1 and L2 reading. In *Interspeech 2013* (pp. 235-238). Lyon, France: ISCA.
- Trofimovich, P., & Baker, W. (2006). Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28, 1-30.
- Ueyama, M. (2000). *Prosodic transfer: An acoustic study of L2 English vs. L2 Japanese* (Doctoral dissertation). University of California. Retrieved from <http://www.linguistics.ucla.edu/faciliti/research/ueyama.pdf>
- Watanabe, K. (1988). Sentence stress perception by Japanese students. *Journal of Phonetics*, 16, 181-186.

Yamato, K. (2004). Realisation of illocutionary force through English intonation by Japanese EFL Learners. *Prospect*, 19, 56-73.

## Appendix

### Printed Transcript Used in the Survey

#### A. Mark the location of "Prominence."

##### 1. Practice Session

they have both their children in a catholic prep school and yeah kindergarten and preschool and it's a yes it's it's a catholic school environment

##### 2. Exercise Session

- 1) she got pregnant and she was like well yknow hey this is my ticket to move out because she gets her own place she gets money for herself money for the baby she doesn't have to pay any utilities or anything
- 2) and as being like the final say in things rather than women and i would see them having an authority figure that's being a woman as causing problems to 'em
- 3) and i did have a very serious relationship in florida uh he was quite a bit younger than i was and i knew he wanted children and i never wanted children unfortunately
- 4) my goodness mom and dad you made a sacrifice because now i'm learning a little more what you did and then by the time you're uh maybe sixty fifty sixty now you got stay care take care of your parents if you have the morals to do it otherwise just ship them off to the farm
- 5) we go to breakfast what'd you do today well let's see i went to breakfast with my mom and then amy and i went to lunch and then we yknow and I went actually

i we moved about four three months ago so i've been kinda doing a lot of stuff at our place so that's nice i mean

- 6) yknow you have these little babies crying and all this stuff going on but he used to always he would come home from work and then go back to work after they were in bed so that i wouldn't have to deal with it the whole time
- 7) um because let's face it to to me love is a choice because there's times when i don't love my husband i wanna smack him and he's being a little barney
- 8) the ymca programs are wonderful they give them a structured activity to do
- 9) or if i had stayed on the east coast where of course it'd be really competitive i i probably wouldn't be working as a freelancer today i came to columbus there wasn't yknow there wasn't much going on when i came here
- 10) but it seems like the one thing that pre almost everybody there agreed on was the didn't wanna see the place go to be over run by condos and housing developments
- 11) give them the money on high street because it really doesn't motivate them to get out of the gutter i mean if they can make a living off of begging change off of college students they're probably gonna stay with that because it's real easy life

#### B. Mark the location of "Chunk boundaries."

##### 1. Practice Session

i have a project i work on not as much as i probably should but it's the swan cleaners project and what we do we have little houses donation boxes in all the swan cleaners



## 2. Exercise Session

- 1) once the tourists hear about it yknow it's like yknow it's like yknow like some of these places in florida kind like ruined by tourists yknow like fort lauderdale Daytona
- 2) and as being like the final say in things rather than women and i would see them having an authority figure that's being a woman as causing problems to 'em
- 3) and uh i could have people rent me a house low income i could move in there i could do all kinds of stuff yknow to make money on the side all through the government
- 4) ya know it's oriented around the church even though we're not catholic it's still oriented as so there's lot of values that permeates everything well they don't do that in the public school because if you do that somebody's gonna get ticked off it's the wrong value and
- 5) yeah i remember going up three fifteen when they first opened it and it was like no body was on it you could just drive on it and it was like an empty road that no body knew about yet and now it's like a parking lot everyday so
- 6) and they have the best reputation of all the columbus schools and uh those kids leave there getting full scholarships uh my son went there one year uh uh it was a disaster for him
- 7) just the suburbs develop and how it's like more of a metropolitan area as opposed to a centralized area that's really big like our like columbus's downtown it's really odd too
- 8) so i wouldn't personally i don't like that myself but i ain't had no problem with them being like a pilot or something like that yknow
- 9) no one was allowed to leave their classroom and then they brought this the dogs through and they found i was surprised they found a lot of drugs
- 10) all the stuff that i mean it was just it was very poorly run to begin with so we would go out and try and find books for the kids i don't know if
- 11) the ymca programs are wonderful they give them a structured activity to do
- 12) they won't take proper advantage of it because the they have their parents to look up to who didn't take advantage of it yknow so if it's kind of the mentality if my parents did okay without it then i don't need it