Talking Points:

The differential impact of real-time computer generated audio/visual feedback on speech-like & non-speech-like vocalizations in low functioning children with ASD

Joshua Hailpern	Karrie Karahalios	Laura DeThorne	Jim Halle
Computer Science	Computer Science	Speech & Hearing Science	Special Education
University of Illinois	University of Illinois	University of Illinois	University of Illinois
201 N Goodwin Ave	201 N Goodwin Ave	901 South Sixth S.	1310 South Sixth St.
Urbana, IL 61802 USA	Urbana, IL 61802 USA	Champaign, IL 61820 USA	Champaign, IL 61820 USA
1-217-333-3328	1-217-265-6841	1-217-333-2230	1-217-244-3557
jhailpe2@cs.uiuc.edu	kkarahal@cs.uiuc.edu	lauras@uiuc.edu	halle@uiuc.edu

ABSTRACT

Real-time computer feedback systems (CFS) have been shown to impact the communication of neurologically typical individuals. Promising new research appears to suggest the same for the vocalization of low functioning children with Autistic Spectrum Disorder (ASD). The distinction between speech-like versus non-speech-like vocalizations has rarely, if ever, been addressed in the HCI community. This distinction is critical as we strive to most effectively and efficiently facilitate speech development in children with ASD, while simultaneously helping decrease vocalizations that do not facilitate positive social interactions. This paper provided an extension of Hailpern et al. in 2009 by examining the influence of a computerized feedback system on both the speech-like and non-speech-like vocalizations of five nonverbal children with ASD. Results were largely positive, in that some form of computerized feedback was able to differentially facilitate speech-like vocalizations relative to nonspeech-like vocalizations in 4 of the 5 children. The main contribution of this work is in highlighting the importance of distinguishing between speech-like versus nonspeech-like vocalizations in the design of feedback systems focused on facilitating speech in similar populations.

Categories and Subject Descriptors

H5.2 [**Information Interfaces and Presentation**]: Screen design, Voice I/O. K4.2 [**Social Issues**]: Assistive technologies for persons with disabilities

General Terms

Experimentation, Human Factors.

Keywords

Accessibility, Visualization, Feedback, Autism, Children, Speech, Vocalization

1. INTRODUCTION

As a child typically develops, vocalization and language skills are acquired with little or no explicit instruction from caregivers. For some children, this skill is not obtained quickly, easily, or at all. One such population is low-functioning individuals with **A**utistic **S**pectrum **D**isorder (ASD). In addition to communication deficits, children with ASD have accompanying difficulties in social interaction and rigidity of behavior. This constellation of impediments not only has detrimental effects on aspects of daily life, but also makes the implementation of many traditional therapies difficult.

One exciting new area of research is directed at creating new technologies to help teach vocalization and potential speech to individuals with ASD. The foundation of this approach is that computers can represent, or reinterpret, the vocalizations of an individual into a new feedback modality (audio or visual). In theory, this could allow an individual to re-experience their vocalizations, gain a new understanding of how they communicate, and potentially adjust how they vocalize [3-5, 14, 15]. Moreover anxiety caused by human-to-human interaction [2, 13] can be reduced through the use of computer-human interaction and thereby increase the subject's willingness to engage. Hailpern et al. [10] conducted a preliminary investigation into this new area of research in 2009, studying the impact of CFS on vocalization in children with ASD.

Like many researchers before them, Hailpern et al. [10] focused exclusively on how CFS impacts speech-like vocalizations. However, in addition to the difficulties in acquiring speech, some children with ASD and other developmental disabilities produce non-speech-like vocalizations that can disrupt successful social interactions [1].

This paper seeks to extend Hailpern et al. [10] by examining the influence of and relationship between **C**omputer **F**eedback **S**ystems (CFS) on both speech-like vocalizations (sounds that can be phonetically transcribed) and non-speech-like vocalizations (e.g. grunt, screech, etc). Ideally, for purposes of speech development, we suggest that the goal should be to increase speech-like vocalizations in the presence of CFS with a decrease in or null effect on non-speech-like vocalizations. It should be noted that the Hailpern et al.'s CFS was not explicitly designed to impact non-speech-like vocalizations (positively or negatively).

By definition, the categories of speech-like and non-speech-like vocalization are mutually exclusive: children can produce one form of vocalization at a time. Consequently, one might anticipate that an increase in one type of vocalization would lead to a decrease in the other. However, within a specified time period, it would be possible to increase the frequency of both types of vocalizations if the initial baseline rate of vocalization is relatively low. It is, therefore, feasible that the CFS may increase the frequency of both speech-like and non-speech-like vocalizations, especially since both types of vocalizations elicited visual and/or auditory feedback from the system.

The main contribution of the present work is to assess the differential impact of CFS on speech-like and non-speech-like vocalizations as well as their interaction and impact on each other. Such information is critical in evaluating the effectiveness of CFS and guiding the development of future feedback systems. A thorough secondary analysis of data previously collected by Hailpern et al.[10] allows us to examine the impact of CFS with a broader set of dependent variables.

We first outline the foundation for these dependent variables as a meaningful form of analysis, as well as situating this work in the existing field of Human Computer Interaction (HCI) and CFS research. We describe the experimental context, our research questions/desired outcome, and the detailed methodology for answering them. We conclude with a discussion of these results both in terms of each question, and at a higher level in terms of the relationship between CFS and vocalization. Beyond the results and data analysis of our data, this work seeks to highlight the need for a multi-dimensional approach to the analysis of CFS and vocalization.

2. RELATED WORK
With an occurrence rate of 1 of every 150 children [7], and approximately half failing to develop functional language [6, 18, 26], the need to effectively advance methods to facilitate the development of communication, including speech, in children with ASD is critical. Over the 60 years since it s initial description [13], scientists have explored many techniques in order to better understand ASD and provide treatments to mitigate its many communicative and social challenges [8, 16, 19, 20, 34]. However, all human-to-human therapeutic techniques share a common drawback: anxiety associated with direct human interaction can make therapy more difficult for many individuals with ASD [2, 13].

Based on research demonstrating the impact of real-time computer based feedback systems on the communication of neurologically typical individuals [3-5, 14, 15], new research has aimed to use computers to help children with ASD [10, 16, 17, 22, 30, 31]. While this work appears quite promising, most studies have examined children who already have well-developed expressive and receptive language skills. Determining if computer systems will be engaging and will elicit vocalizations by individuals without basic communication is a largely unaddressed question.

One notable exception is Hailpern et al. in 2009, who performed a pilot study examining the effect of CFS on the spontaneous speech-like vocalizations of low functioning children with ASD. One limiting aspect of this work was the exclusive focus on speech-like vocalizations. There is a wealth of descriptive literature that highlights the differences in speech-like and non-

speech-like vocalizations [28, 29, 32, 33], and we decided to expand the analysis of Hailpern et al. [10] accordingly; examining both speech-like and non-speech-like vocalization. Conceptually, these two types of vocalizations are mutually exclusive. An individual can only produce one vocalization type at a time. Further, there may also be a relationship between how often someone vocalizes and how long each vocalization lasts. Therefore, when designing solutions to encourage vocalization, we must consider the impact of CFS as potentially multi-faceted. There are very few intervention studies that have looked at or targeted their impact on both speech-like and non-speech-like vocalizations ([1] as one notable exception). As a result, we aim to uncover the impact of CFS on vocalization in terms of speech-like and non-speech-like vocalization, while simultaneously evaluating the potential relationship between the two.

3. EXPERIMENTAL CONTEXT
To examine the effect of CFS on both speech-like and non-speech vocalizations in low functioning children with Autistic Spectrum Disorder, we have re-examined the data from *The Spoken Impact Project* experiment [10]. The CFS used by Hailpern et al. [10] was not explicitly designed to impact non-speech-like vocalizations, and thus was not examined in their initial analysis. Using the raw data, we re-analyzed the trials in each session in order to assess the impact of CFS using both speech-like and non-speech-like vocalizations as dependent variables.

4. QUESTIONS
Our analyses are organized around four specific questions that tease apart the relationship between CFS, speech-like vocalizations, and non-speech-like vocalizations. This section details each question, and outlines our desired outcomes. Measurements are centered on *Speech-Like Vocalizations* (SLV) and *Non-speech-Like Vocalizations* (NSLV). Section 5.3.4 provides further details on and definitions of SLV and NSLV. Section 5.2 describes the CFS used.

4.1 Q1: Does CFS Impact Children's Frequency and/or Duration of Vocalization?
Question 1 examines, at the most general level, the impact of CFS on vocalizations in low functioning children with ASD. To conduct this analysis we compared baseline trials (no feedback) to all other trials (regardless of feedback style or modality). We assessed the impact on three dependent variables: rate of SLV, average duration of SLV, and rate of NSLV. Because our data does not contain NSLV duration, we were unable to assess the impact on NSLV duration. Upon further consideration, this may have been a difficult variable to assess in that many NSLVs are nearly instantaneous (e.g. lip pop), and thus difficult or impractical to measure. *Note that a subset of SLV (those vocalizations that were not imitative) was the only dependent variable examined in the paper by Hailpern et al* [10].

4.1.1 Desired Outcome
Our desired outcome is that CFS would increase rate and/or duration of SLVs while decreasing the rate of NSLVs. In other words, we hope that CFS would encourage children to vocalize in a manner that could be used for speech, while discouraging non-speech vocalizations (e.g. screaming, grunts, etc) that may be viewed as less socially acceptable. This is based on how CFS was introduced to each child; specifically the CFS was demonstrated

for each child through use of speech. Specifically, the investigators would draw each child's attention to the screen while saying words such as the child's name or an exclamation (e.g., "Boo!"). Although imitation is an area of difficulty for many children with ASD [12], this does not mean the skill is nonexistent. Imitation is a commonly employed strategy used by children and adults in unfamiliar situations.

4.2 Q2: Does CFS Differentially Increase SLV Relative to NSLV?

Our second question examined potential trade-offs between rate of SLV, and rate of NSLV. By examining the overall change in vocalizations from *both* sides of the vocalization spectrum, we can best examine the overall impact of CFS. This analysis will follow Q1 by comparing baseline to all trials, while examining the effect of CFS through a different dependent variable, the ratio of SLV to NSLV:

> *SLV:NSLV Ratio – This dependent variable measures the relative change between potentially phonetically transcribable vocalizations and those that are not.*

An increase in one form of vocalization may lead to a decrease in another form or both types may increase as an expression of decreased silence.

4.2.1 Desired Outcome
Consistent with the rationale we presented in relation to Q1, our desired outcome is less frequent NSLV compared to SLVs, indicating an increase of SLV compared to NSLV during CFS trials. Thus our ratio should get larger in non-baseline conditions.

4.3 Q3: Are the different vocalization measures correlated?

To better understand if there is a significant relationship between SLV rate, SLV duration and/or NSLV rate, we conducted correlations between the different dependent variables. We examined the relationship between dependent variables within each child.

4.3.1 Desired Outcome
We hope that there will be a negative correlation between speech-like vocalizations and non-speech-like vocalizations indicating that in the presence of one, the other decreases.

4.4 Q4: How does Modality Impact Vocalizations

To examine the differences between different modalities of feedback (audio/visual), we sought to re-examine Q1 and Q2 while comparing baseline to conditions with *Audio Only* (no visual feedback), *Visual Only* (no audio), and *Mixed* (both visual *and* auditory feedback must be present).

- *Does the impact of CFS on children's vocalizations (SLV and NSLV) differ as a function of feedback modality (audio, visual, mixed)?*
- *Does the impact of CFS on children's SLV:NSLV ratio differ as a function of feedback modality (audio, visual, mixed)?*

	Oliver	Frank	Larry	Diana	Brian
Age (years)	5	8	4	4	3
Diagnosis	ASD	ASD + Down's syndrome	ASD + Down's syndrome	ASD	ASD

Table 1. Demographics of participants.
Names changed to protect identity, but gender status was retained.

4.4.1 Desired Outcome
We hope that the impact of CFS would differ as a function of modality. Though the *common thought* is that children with ASD respond more favorably to visual stimuli, based on the previous findings of Hailpern et al., we anticipated both audio and visual feedback would produce effects that would differ for individual children.

5. METHODS
Outlined here are the basic experimental design and methods. However, we encourage the reader to refer to the original paper detailing the *Spoken Impact Project* for full experimental and software design specifications [10]. The original experiment was conducted over approximately six months during 2007.

5.1 Participants
In order to examine the effects of CFS on low functioning children with ASD, Hailpern et al. recruited five children whose ages ranged from 3 to 8 years (see Table 1). All five of the subjects' spoken language developmental benchmarks [24] were in the first stage (Preverbal Communication), roughly comparable to the development of a neurologically typical 6-12 month old. To protect the privacy of the subjects, individual's names were changed; gender status was retained.

5.2 Computer Feedback System
The CFS used was a collection of audio and visual feedback that translated a child's vocalizations into a new medium (audio and visual respectively). The CFS used was named *The Spoken Impact Project Software (SIPS)*. This real-time computer feedback system detected changes in volume, vocalization duration, pause, and pitch and presented feedback mapped to these parameters. For example, consider a circle in the center of a computer screen, whose diameter increases as volume increases, and its color changes after each two-second pause. Similarly, for audio feedback, the computer provided an echo of the subject's voice (directly matching their vocalizations) or provided a personal pre-recorded sound clip or musical sound as an audio reward (whose duration was directly proportional to the duration of subject's vocalization) at the completion of each utterance. A detailed explanation of feedback styles and modalities is found in [10].

5.3 Procedures
5.3.1 Adapted Within-Subject Design
Due to the relatively unexplored nature of this task in this population, Hailpern et al.'s investigation was a preliminary pilot study of five subjects. This experimental design permitted an examination of each subject independent of the others. That is, each subject served as his or her own control and followed the

dictates of a within-subject investigation. Each subject had a baseline trial against which the feedback conditions were evaluated.

5.3.2 Experiment/Session/Trial Design

Each of the five subjects participated in six sessions. A session lasted approximately 40 minutes and was divided into approximately eight 2-minute trials. The first trial of each session was a baseline trial during which no feedback (a black screen with no sound) was delivered. This trial functioned as a control and permitted a comparison for experimental trials that consisted of exposure to one permutation of audio and/or video CFS. Following each trial, subjects were given a two-minute break period consisting of preferred food/toys/games. This was designed to encourage a positive and friendly atmosphere. Trial length and frequency varied slightly depending on child cooperation.

5.3.3 Independent Variables

Trials included varying forms of visual and auditory CFS. Subsequent to baseline trials, each trial included at least one form of visual or auditory feedback. Some trials included both auditory and visual feedback (*Mixed*).

5.3.4 Dependent Variables

The specific dependent variables used in each analysis were described in Section 4. Each dependent variable is a composite of one of two measures, *Speech-Like Vocalizations* (SLV) and *Non-speech-Like Vocalizations* (NSLV). We calculated SLV and NSLV in terms of rate rather than frequency to insure and equivalent metric across trials that differed in length. We define rate as frequency of vocalization during a trial divided by trial duration (freq/time). All time-based measurements were made in milliseconds. Both SLV and NSLV are drawn from the A^3 coding guidelines published in 2008 [11].

Speech-Like Vocalizations– sounds produced by the subject that can be phonetically transcribed based on spoken English phonology. SLV was measured both in terms of rate and duration. Both measures were considered given the potential trade-off between rate and duration during a fixed time period. Longer SLV could translate into a lower rate of SLV.

Non-speech-Like Vocalizations- vocalizations that cannot be phonetically transcribed. These are sounds that relate less directly to conventional speech. Examples include ticks, heavy breathing, vocal stereotypes, or other forms of expression that are not used in speech production (laughing, screaming, etc) [11, 34].

Throughout the remainder of the paper, we use a postfix of **–R** to denote variables of rate and **–D** to denote variables of duration.

The analysis conducted in Hailpern et al. [10] exclusively examined vocalization in terms of *Spontaneous Speech-Like Vocalization Rate* (a subset of SLV; specifically, those speech-like vocalizations that are not imitative [11]). We included both spontaneous and imitative SLV here to maximize power. *Note this may lead to minor differences in results from the initial analysis in [10].*

5.3.5 Data Collection

Data were gathered during a six-month period representing a 1200-minute corpus of video. The video was annotated by two coders, both with BS degrees in Speech and Hearing Science (one pursuing an MA). Video was annotated using the VCode and VData system [9]. The **i**nter-**r**ater **a**greement (IRA) was calculated using point-by-point agreement, resulting in an overall IRA of 88%. SLV frequency had an IRA of 91%, and durational IRA of 93%. Further, NSLV frequency had an IRA of 84%.

	SLV-R (%)	NSLV-R (%)	SLV-D (%)
Maximum Power	100	99	96
Minimum Power	5	6	9
Mean Power	35	41	34
Median Power	12	35	26

Table 2. Statistical Power By Measure
Power calculated across subjects, and across all sessions/trials.

5.3.6 Statistical Tests

Given the non-parametric nature of the data collected, traditional paired t-tests would not be applicable. The Wilcoxon Rank-Sum was used as a non-parametric alternative. Similarly, the Spearman rank correlation test was used as a non-parametric alternative to the Pearson correlation test.

5.3.7 Significance

All tests used a two-tailed alpha with a p<.05 to denote statistical significance. We indicate statistical significance with an asterisk (*). While performing multiple comparisons may suggest statistical adjustment to a more conservative value (i.e., Bonferroni correction), the small number of data points and low statistical power of our study would suggest a more relaxed threshold (See Table 2). Taking both perspectives into consideration, an alpha of <.05 was determined to be a reasonable approach. Moreover, as a preliminary pilot study, the focus of this research was to highlight avenues of future research rather than to

	NSLV-R			SLV-R			SLV-D			SLV-R/NSLV-R		
Oliver	2.24,	*0.025**	↓	-1.92,	0.055		1.17,	0.240		-2.82,	*0.005*†*	↑
Frank	-2.38,	*0.017**	↑	-2.51,	*0.012*†*	↑	1.07,	0.286		-1.66,	0.098	
Larry	0.62,	0.535		-0.22,	0.829		1.99,	*0.046**	↓	-0.18,	0.861	
Diana	0.99,	0.324		-0.10,	0.920		-0.27,	0.789		-1.32,	0.188	
Brian	2.08,	*0.038**	↓	0.12,	0.907		1.54,	0.124		-2.03,	*0.042**	↑

Table 3. Wilcoxon Rank-Sum across all trials. Rate (-R) & Duration (-D) of SLV and NSLV. (z value, p value)
arrow represents statistically significant change/direction in measure/condition vs. baseline
* denotes significance < .05, † denotes significance in Sensitivity Analysis (See 5.3.7)

test competing hypotheses. Consequently, following the rationale of Savitz [27], we believe that the potential of detecting false positives (5%) was outweighed by the concern of missing meaningful effects. *For all results and discussion, a p<.05 will be used to denote significance.*

Nonetheless, we performed a sensitivity analysis in which p-values were adjusted using a Bonferroni correction for multiple rank-sum comparisons. For the dependent variable SLV-D, we performed two statistical comparisons on each independent data set, resulting in an adjusted threshold of .0250 for statistical significance. Similarly, the SLV-R and NSLV-R variables were used in four statistical comparisons on each independent data set, resulting in a threshold of .0125 for statistical significance. We indicate statistical significance for the sensitivity analysis with a dagger (†) in Table 3 and Table 5.

6. RESULTS

We present here the results related to each research question. We present a summary of the results in addition to the statistical findings. Due to the within-subject design of this experiment, we will further present the results for each child independently. Please refer to Table 1 for subject's pseudo name, age and diagnosis (consistent with [10]). A detailed discussion of the results and their implications is presented in the following section.

6.1 Q1: Does CFS Impact Children's Rate and/or Duration of Vocalization?

CFS appeared to influence the rate or duration of vocalizations in four of the five children. (See Table 3)

With all feedback types collapsed, **Oliver** significantly decreased NSLV-R while our analysis trended towards increasing his SLV-R with *p=0.055*. **Frank** significantly increased both SLV-R and NSLV-R in conditions with CFS. **Larry** significantly *decreased* his SLV-D during CFS. **Diana** showed no change in NSLV-R, SLV-R, or SLV-D. **Brian** significantly decreased NSLV-R with CFS. Consequently, when all feedback trials were combined desirable effects of either increasing SLV or decreasing NSLV were observed in 3 out of 5 children with only one child showing an undesirable decrease in SLV-D.

6.2 Q2: Does CFS Differentially Increase SLV Relative to NSLV?

When the overall impact of the CFS is considered, only two out of five children increased their ratio of SLV-R relative to NSLV-R, but primarily by decreasing the freq of NSLV. (See Table 3)

Oliver significantly increased SLV-R:NSLV-R ratio through a decrease of NSLV-R. **Frank** had no significant change when comparing SLV-R to NSLV-R. It is of interest, however, that Frank increased both his SLV-R and NSLV-R, resulting in a net ratio gain. **Larry** and **Diana** presented no significant change in relative use of SLV-R to NSLV-R, at least when results from all feedback types were collapsed. **Brian** significantly increased his SLV-R per NSLV-R primarily through a decrease in NSLV-R during CFS. *Note that findings from the omnibus test may mask important individual differences in children's response to particular modalities/forms/styles of feedback as shown in* [10].

	SLV-D by SLV-R		SLV-R by NSLV-R		NSLV-R by SLV-D	
Oliver	**-.451,**	**<.001***	-.049,	.710	-.198,	.138
Frank	**-.430,**	**.001***	**-.399,**	**.002***	.027,	.846
Larry	**.278,**	**.048***	.208,	.135	.099,	.488
Diana	.179,	.249	-.168,	.282	.208,	.182
Brian	.125,	.429	**.334,**	**.027***	-.122,	.440

Table 4. Spearman Rank Correlations (rho, p)
rho indicates relative slope between variables, and effect size
* denotes significance < .05

6.3 Q3: Are the different vocalization measures correlated?

When considering the correlations among the three dependent variables, we noted significant relationships between SLV-D and SLV-R for three of the five subjects and significant relationships between SLV-R and NSLV-R for two out of five subjects. (See Table 4)

Oliver and **Frank** show a negative correlation between SLV-D and SLV-R, while **Larry**'s correlation was positive. Further, both **Frank** and **Brian**'s SLV-R and NSLV-R correlations were significant, however, the direction of the association was in opposite directions from each other.

6.4 Q4: How does Modality Impact Vocalizations

When considering the effect of feedback modality on vocalization, each child appeared to have a unique response to different forms of feedback. Specifically, two of the five subjects had a significant response to only audio feedback; three of the five subjects had a significant response to conditions with only visual feedback, and all five subjects responded significantly to the mixed feedback condition. The following subsections and Table 5 detail the responses of each child.

Consistent with the negative correlation in Table 4, **Oliver** significantly increased his SLV-R and decreased his SLV-D in conditions with only audio feedback. In addition, Oliver's SLV-R to NSLV-R ratio increased in both the auditory only and mixed conditions. This suggests that for Oliver audio feedback was a useful form of CFS. While the presence of visual feedback with the audio did not change his SLV-R or NSLV-R, it did improve the ratio between the two.

Frank significantly increased his NSLV-R in all conditions with visual feedback (visual only and mixed). The absence of change in NSLV with only audio suggests that visual feedback was the primary catalyst for increasing NSLV-R. Further, Frank increased his SLV-R in the conditions with only audio feedback. The absence of an increase in SLV-R in conditions with visual feedback suggests that Frank responded favorably to audio feedback and unfavorably to visual feedback in terms of desired responses. Additional support for this conclusion comes from the positive increase in Frank's SLV-R to NSLV-R ratio in conditions with audio feedback (audio only and mixed) paired with the decreased ratio in conditions with visual only feedback. Together these results suggest a cohesive picture for Frank with audio

		NSLV-R			SLV-R			SLV-D			SLV-R:NSLV-R		
Visual Only	Oliver	1.52,	0.129		-0.87,	0.386		-0.82,	0.409		-1.68,	0.093	
	Frank	-2.95,	*0.003*†*	↑	0.47,	0.637		-1.29,	0.196		2.01,	*0.044**	↓
	Larry	-1.29,	0.197		0.26,	0.796		0.33,	0.739		1.29,	0.197	
	Diana	-0.10,	0.920		0.90,	0.366		-2.41,	*0.016*†*	↑	-0.40,	0.692	
	Brian	2.09,	*0.037**	↓	0.19,	0.850		1.34,	0.182		-2.04,	*0.041**	↑
Audio Only	Oliver	1.57,	0.116		-2.14,	*0.032**	↑	2.00,	*0.046**	↓	-2.72,	*0.007*†*	↑
	Frank	-0.85,	0.394		-2.56,	*0.011*†*	↑	0.00,	1.000		-2.57,	*0.010*†*	↑
	Larry	1.32,	0.187		0.16,	0.869		1.57,	0.117		-0.54,	0.591	
	Diana§												
	Brian§												
Mixed Feedback	Oliver	-1.93,	0.054		-0.33,	0.742		1.40,	0.161		-2.74,	*0.006*†*	↑
	Frank	-3.00,	*0.003*†*	↑	1.70,	0.089		1.70,	0.089		-2.32,	*0.021**	↑
	Larry	-0.44,	0.660		1.03,	0.304		2.10,	*0.036**	↓	-0.20,	0.839	
	Diana	-0.54,	0.588		-2.14,	*0.032**	↑	0.72,	0.470		-1.64,	0.100	
	Brian	-0.26,	0.796		1.64,	0.101		2.24,	*0.025*†*	↓	-1.51,	0.131	

Table 5. Wilcoxon Rank-Sum across modalities. Rate (-R) & Duration (-D) of SLV and NSLV. (z value, p value)
arrow represents statistically significant change/direction in measure/condition vs. baseline
* denotes significance < .05, † denotes significance in Sensitivity Analysis (See 5.3.7)
§audio only feedback was not tested [10] due to lack of interest observed in initial orientation sections

feedback serving as the primary catalyst for SLV, and visual feedback as the primary catalyst for NSLV.

Larry significantly decreased his SLV-D in the mixed feedback condition. Given the lack of significance found in visual only and audio only conditions, the data suggest that Larry's decrease in SLV-D is primarily due to the presence of audio *and* visual feedback together. Overall, Larry does not appear to be a good candidate for this particular CFS based therapy.

Diana significantly increased her SLV-D in conditions with only visual feedback, and significantly increased her SLV-R in the presence of mixed feedback. This indicates that Diana responded primarily to visual feedback, in particular a specific type of visual feedback (see [10] for details).

Brian decreased his NSLV-R in conditions with only visual feedback. This indicates that visual feedback may have suppressed his NSLV. Consistently, Brian's SLV-R:NSLV-R ratio increased during conditions with visual-only feedback.

	NSLV-R		SLV-R		SLV-D		SLV-R :NSLV-R	
	↑	↓	↑	↓	↑	↓	↑	↓
Visual Only	F	B	-	-	D	-	B	F
Audio Only	-	-	OF	-	-	O	OF	-
Mixed	F	-	D	-	-	LB	OF	-
# of Unique Subjects	1	1	3	0	1	3	3	1

Table 6. Unique Subjects With Significant Changes Compared to Baseline in Different Conditions
First initial of subject's name represents change in behavior relative to baseline (arrows represent change direction)

While not enough data were collected to determine the effect of audio-only feedback, the lack of a statistically significant response in all other conditions does suggest that audio feedback did not play a role in improving his ratio.

7. DISCUSSION

Examination of changes in the SLV-R:NSLV-R ratio in Table 5 indicates that for three of five of subjects adjusted the ratio changed in the presence of CFS. However, for the five conditions in which there was an improved ratio (more SLV-R per NSLV-R), one was due to suppress NSLV-R and two were due to increased SLV-R.

Our data suggest that CFS had a significant impact on all five subjects for at least one measure. However, the specific outcome varied by subject and by modality. This indicates that individual differences in subjects' response to computerized feedback should be considered in any attempt to develop feedback for the purpose of facilitating speech. In particular, the inclusion of audio as a significant modality is noteworthy here as the "conventional wisdom" is that visual feedback is particularly engaging to individuals with ASD [21, 23, 25]. Although highly preliminary in nature, findings suggest that audio feedback should not be automatically discounted as a means to facilitate speech development, particularly if the audio feedback is shaped in a way that is not available in everyday encounters.

Table 6 illustrates the breakdown of the number of subjects whose vocalizations significantly changed with CFS. From this high level analysis, we are brought back to our basic questions regarding suppression vs. encouragement of SLV and NSLV, as well as the relationship between the two variables. Of particular note, no feedback conditions appeared to suppress SLV-R, which is critical to CFS being useful in a therapeutic condition.

Further, the data suggests that in certain conditions there may be a trade-off between SLV rate and SLV duration. When we examine the effect of CFS on SLV-D (particularly in Table 6), an effect of suppression appears, rather than encouragement. When considered together increases in SLV-R, evidence of this trade-off appears. This is further supported by examination of Table 4 and the negative correlations of Oliver and Frank (the exception being Larry with a positive correlation). In other words, there may be a trade-off between rate and duration of SLV: as subjects vocalize more, the individual vocalizations may become shorter. Consequently, thought should be given, both by investigators and by clinicians, as to the specific goals of their intervention, and recognize that some outcomes may in fact compete with one another. For researchers moving forward with similar work, this finding suggests that it may be naive to only examine rate or duration in isolation. Rather, an examination of both may be necessary so as not to miss the potentially important underlying effects of their CFS.

In addition to the potential for differentiated effects on vocalization rate and duration, recognizing the potential to differentially impact speech-like versus nonspeech-like vocalizations is key. Feedback appeared to be effective in suppressing NSLV-R relative to SLV-R. This result suggests that the vocalizations elicited by CFS may be more adaptive for teaching spoken language skills. It is hard to fully understand the source of this relative change given the limited data of the study, which leads us to caution that the effects of any particular feedback on vocalization should not be assumed to be positive. The differential influence of visual versus audio feedback for Frank highlights this point, with the former facilitating nonspeech-like vocalizations and the latter speech-like vocalizations. That being said, CFS generally appeared to affect SLV-R more so than NSLV-R across the five children (see Table 6). Such results provide hope that CFS can encourage increased SLV-R vocalizations while decreasing or having minimal impact on NSLV-R.

Given our focus on statistical significance, we want to explicitly note that the absence of significance in a specific condition/modality does not indicate lack of potential impact. In [10], it was shown that specific feedback modalities and individual styles within that modality were often most successful for individual children. This finding was similarly noted in the qualitative analysis, particularly in regard to specific modalities. Yet Hailpern et al.'s study did *not* design or account for subject preference. In other words, the ideal form of feedback for each child in each modality may not have been explicitly built into the system. While our results do suggest conditions where CFS had impact, researchers should not rule out any of our conditions that did not have significant findings; the nature of this particular CFS may not have been engaging to that individual child.

8. LIMITATIONS

Due to the small scale and low statistical power of this study, conclusions drawn are at best preliminary. However, the purpose of both Hailpern et al.'s original study and this secondary analysis are not to provide definitive conclusions on the impact of CFS, but rather to suggest its potential as a new direction of HCI research and to highlight factors that researchers should take into consideration. Because of multiple testing and our decision not to make a statistical adjustment, it is possible that some of our significant findings may have been spurious. While these

limitations are important to note, they do not greatly hamper the position or potential of the findings to shape future investigations.

9. CONCLUSION

Given the findings of this data analysis, we believe that Computer Feedback Systems (CFS) present an exciting and potentially promising new technique for encouraging vocalization in low functioning children with ASD. More specifically, CFS appears to encourage the rate of vocalizations that can be considered *Speech-Like* (sounds that can be phonetically described), while generally having little impact on those vocalizations that we call *Non-speech-Like* (e.g. grunt, screech, etc). In other words, CFS has the potential to differentially impact (favorably) speech-like versus non-speech-like vocalizations. However, the nature of this potential relationship needs to be directly considered rather than implicitly assumed for any particular child, form or modality of feedback.

This understanding strengthens the potential impact of CFS, and the original study [10] from which this data was collected. Together, these studies help delineate the need for future research in computerized feedback systems in order to better understand its impact on vocalization and potential in helping teach language and communication skills to children with Autistic Spectrum Disorder.

10. ACKNOWLEDGMENTS

11. REFERENCES

[1] Baas, B. S., Strand, E. A., Elmer, L. M. and Barbaresi, W. J. Treatment of Severe Childhood Apraxia of Speech in a 12-Year-Old Male with CHARGE Association. *Journal of Medical Speech-Language Pathology*, 16, 4 2008), 9.

[2] Baskett, C. B. *The effect of live interactive video on the communicative behavior in children with autism*. University of North Carolina at Chapel Hill, Chapel Hill, 1996.

[3] Bergstrom, T. and Karahalios, K. Conversation Clock: Visualizing Audio Patterns in Co-Located Groups. In *Proceedings of the HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCE 2007* (Hawaii, USA, 2007). IEEE, 2007.

[4] Bergstrom, T. and Karahalios, K. Conversation votes: enabling anonymous cues. In *Proceedings of the CHI '07 extended abstracts on Human factors in computing systems* (San Jose, CA, USA, 2007). ACM, 2007.

[5] Bergstrom, T. and Karahalios, K. Seeing More: Visualizing Audio Cues. In *Proceedings of the INTERACT* (Rio de Janerio, Brasil, 2007). ACM Press., New York, NY, 2007.

[6] Bosseler, A. and Massaro, D. Development and evaluation of a computer-animated tutor for vocabulary and language learning in children with autism. *Journal of Autism and Developmental Disorders*, 33, 6 2003), 653-672.

[7] Center for Disease Control and Prevention, C. Autism Information Center, DD, NCBDDD, CDC. http://www.cdc.gov/ncbddd/autism/ Atlanta, Accessed 2007

[8] Greenspan, S. I. and Wieder, S. Developmental Patterns and Outcomes in Infants and Children with Disorders in Relating and Communicating: A Chart Review of 200 Cases of Children with Autistic Spectrum Diagnoses *The Journal of Developmental and Learning Disorders* 1, 1 1997).

[9] Hagedorn, J., Hailpern, J. and Karahalios, K. G. VCode and VData: Illustrating a new Framework for Supporting the Video Annotation Workflow. In *Proceedings of the Advanced Visual Interfaces* (Napoli, Italy, 2008). ACM-PRESS, New York, NY, 2008.

[10] Hailpern, J., Karahalios, K. and Halle, J. Creating a Spoken Impact: encouraging vocalization through audio visual feedback in children with ASD. In *Proceedings of the CHI 2009* (Boston, MA, 2009). ACM, 2009.

[11] Hailpern, J., Karahalios, K., Halle, J., DeThorne, L. S. and Coletto, M. A3: A Coding Guideline for HCI+Autism Research using Video Annotation. In *Proceedings of the ACM SIGACCESS- ASSETS 2008* (Halifax, Canada, 2008). ACM-PRESS, New York, NY, 2008.

[12] Jones, V. and Prior, M. Motor imitation abilities and neurological signs in autistic children. *Journal of Autism and Developmental Disorders*, 15, 1 1985), 37-46.

[13] Kanner, L. Autistic Disturbances of Affective Contact. V.H. Winston, 1943.

[14] Karahalios, K. and Donath, J. Telemurals: linking remote spaces with social catalysts. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (Vienna, Austria, 2004). ACM New York, NY, USA, 2004.

[15] Karahalios, K. G. and Dobson, K. Chit chat club: bridging virtual and physical space for social interaction. In *Proceedings of the CHI '05 extended abstracts on Human factors in computing systems* (Portland, OR, USA, 2005). ACM, 2005.

[16] Koegel, L. K., Koegel, R. L., Harrower, J. K. and Carter, C. M. Pivotal Response Intervention I: Overview of Approach. *The Journal of The Association for Persons with Severe Handicaps*, 24, 3 (Fall 1999), 174-185.

[17] Lehman, J. F. Toward the use of speech and natural language technology in intervention for a language-disordered population. In *Proceedings of the Proceedings of the third international ACM conference on Assistive technologies* (Marina del Rey, California, United States, 1998). ACM Press, New York, NY, 1998.

[18] Lord, C., Rutter, M. and Couteur, A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 5 1994), 659-685.

[19] Lovaas, I. I. *The Autistic Child.* John Wiley & Sons, Inc, New York, 1977.

[20] Lovaas, O. I. *Teaching Individuals with Developmental Delays: Basic Intervention Techniques.* PRO-ED, Inc., Austin, TX, 2003.

[21] Minshew, N. J., Goldstein, G. and Siegel, D. J. Neuropsychologic Functioning in Autism: Profile of a complex information processing disorder. *Journal of the International Neuropsychological Society*, 3. 1997), 303-316.

[22] Mohamed, A. O., Courboulay, V., Sehaba, K. and Menard, M. Attention analysis in interactive software for children with autism. In *Proceedings of the ACM SIGACCESS Conference on Assistive Technologies* (Portland, Oregon, October 23 - 25, 2006). ACM, 2006.

[23] Ozonoff, S., Pennington, B. and Rogers, S. Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry*, 32, 7 1991), 1081-1105.

[24] Paul, R. Recommendation for Benchmarks - Autism Speaks Luncheon at SRCLD. L. DeThorne. 2008.

[25] Peterson, S., Bondy, A., Vincent, Y. and Finnegan, C. Effects of altering communicative input for students with autism and no speech: Two case studies. *Augmentative and Alternative Communication*, 11, 2 1995), 93-100.

[26] Prizant, B. Language Acquisition and Communicative Behavior in Autism Toward an Understanding of the" Whole" of It. Journal of Speech and Hearing Disorders, volume 48, Aug 1983), 296-307.

[27] Savitz, D. A. and Olshan, A. Multiple Comparisons and Related Issues in the Interpretation of Epidemiologic Data. *American Journal of Epidemiology*, 142, 9 1995), 4.

[28] Sheinkopf, S. J., Mundy, P., Oller, D. K. and Steffens, M. Vocal Atypicalities of Preverbal Autistic Children. *Journal of Autism and Developmental Disorders*, 30, 4 (August 2000), 345-354.

[29] Stone, W. and Caro-Martinez, L. Naturalistic Observations of Spontaneous Communication in Autistic Children. *Journal of Autism and Developmental Disorders*, 20, 4 1990), 16.

[30] Tartaro, A. Storytelling with a virtual peer as an intervention for children with autism. *SIGACCESS Access. Comput.*, 84 2006), 42-44.

[31] Tartaro, A. and Cassell, J. Playing with Virtual Peers: Bootstrapping Contingent Discourse in Children with Autism. In *Proceedings of the Proceedings of International Conference of the Learning Sciences* (Utrecht, Netherlands, June 24-28, 2008). ACM Press, 2008.

[32] Wetherby, A. and Prutting, C. Profiles of Communicative and Cognitive-Social Abilities in Autistic Children. *Journal of Speech and Hearing Research*, 27, September 1984), 13.

[33] Wetherby, A. M., Yonclas, D. G. and Bryan, A. A. Communicative Profiles of Preschool Children with Handicaps: Implications for Early Identification. *Journal of Speech and Hearing Disorders*, 54, May 1989), 10.

[34] Woods, J. J. and Wetherby, A. M. Early Identification of and Intervention for Infants and Toddlers Who Are at Risk for Autism Spectrum Disorder. *Language, Speech, and Hearing Services in Schools*, 34, July 2003 (July 2003), 180-193.