The University of Cincinnati
College of Allied Health
Department of Communication Sciences and Disorders

Automatic Syllabic Cluster Analysis of Children's Speech Data to Identify Speech-Disorders

Marisha Speights, Keiko Ishikawa, Suzanne Boyce, Joel MacAuslan,
Harriet Fell, JoHannah Ungruhe, & Kailey Longpre

Acknowledgements

Speech Technology and Research Corp. (STAR)

NIH SBIR Grant 5R44DC010104-04

NIH Pre-doctoral Grant to Promote Diversity in Health-Related Research (3R44DC010104-03S1)

SpeechMark®

NIH

National Institutes of Health

UNIVERSITY OF Cincinnati

Introduction

- This research investigates syllabic complexity in children with normal and disordered speech production using a computerized method of analysis.
- Automatic Syllabic Cluster Analysis based upon landmark theory (Stevens 1992, 2002; Liu 1996; Howitt 2000; Fell & MacAuslan, 2005) is used to automate the analysis of child speech
- The algorithm automatically measures acoustic changes that correspond to syllable patterns and provides a fast method for measuring complexity in syllable production without the need for phonetic transcription.

UNIVERSITY OF Cincinnati

Background of Study

Speech development in children involves:

- (a) increasing the proportion of multisyllabic words produced
- (b) moving beyond one or two syllable types (V, VC, CVC) to a larger number of complex syllable types (CCVC, CVCC, CCVCC, etc. (Oller et al., 1999, Oller, 2000)
- The development of well-formed syllables in infancy has been shown to be a significant predictor of later communication skills (Oller et al., 1999, 2010; Oller, 2000; Nathani et al., 2006; Pharr et.al,2000)

UNIVERSITY OF Cincinnati

The Problem of Measuring Speech Complexity

- No universally accepted definition of complexity.

- Systematic guidelines for evaluating complexity in continuous speech samples are not well established.

- Conventional methods of analyzing syllable, word, or utterance complexity are slow and laborious (i.e. phonetic transcription and hand scoring of speech data.

The Problem of Phonetic Transcription

- Even Typically-Developing Children are variable and imprecise in their speech patterns.

- Transcribers tend to “regularize” this variability, as they naturally tend to attempt “make sense” out of an utterance.

- Transcription of large data sets is time- and labor-intensive, thus restricting ability to study large boluses of naturalistic speech (Oller, 2010).

UNIVERSITY OF Cincinnati

Automatic Syllabic Cluster Analysis: Landmark Analysis

- Landmark analysis is based on the work of Stevens et al (2002).

- Aims to identify points in the acoustic signal that are most perceptually salient for information about phonemes, words and meaning.

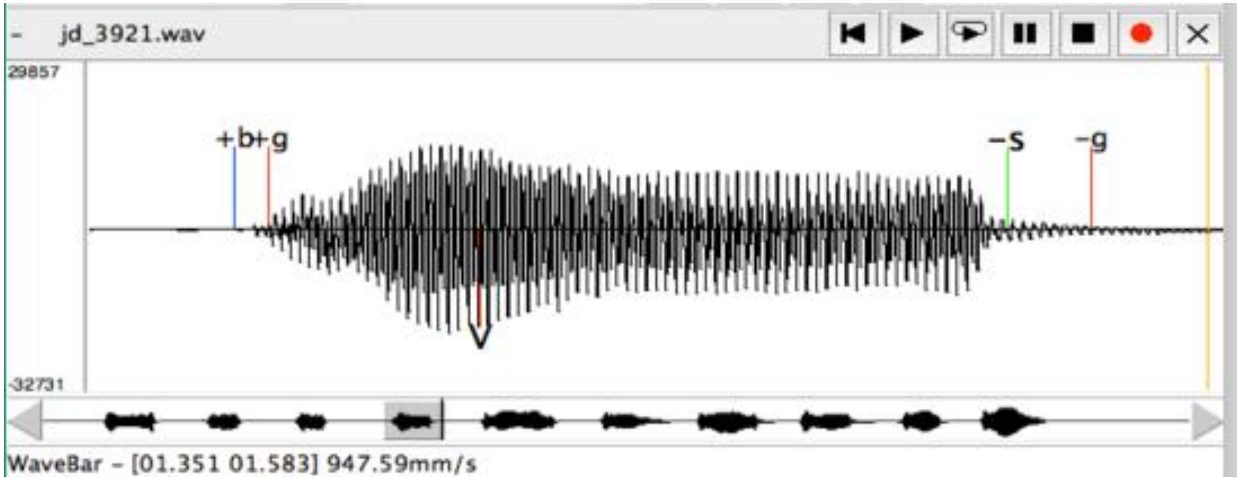
- Looks for patterns of **abrupt change** and **maxima/minima** that occur **simultaneously** across a wide range of frequencies

UNIVERSITY OF Cincinnati

Syllabic Cluster Analysis

– The Syllabic Cluster algorithm in the SpeechMark® Matlab toolbox uses six landmark types and groups sequences of landmarks into syllabic clusters.

– The six abrupt landmarks used are onset and offset versions of +/-**g** (glottal), +/-**b** (noise burst), and +/-**s** (sonorancy).



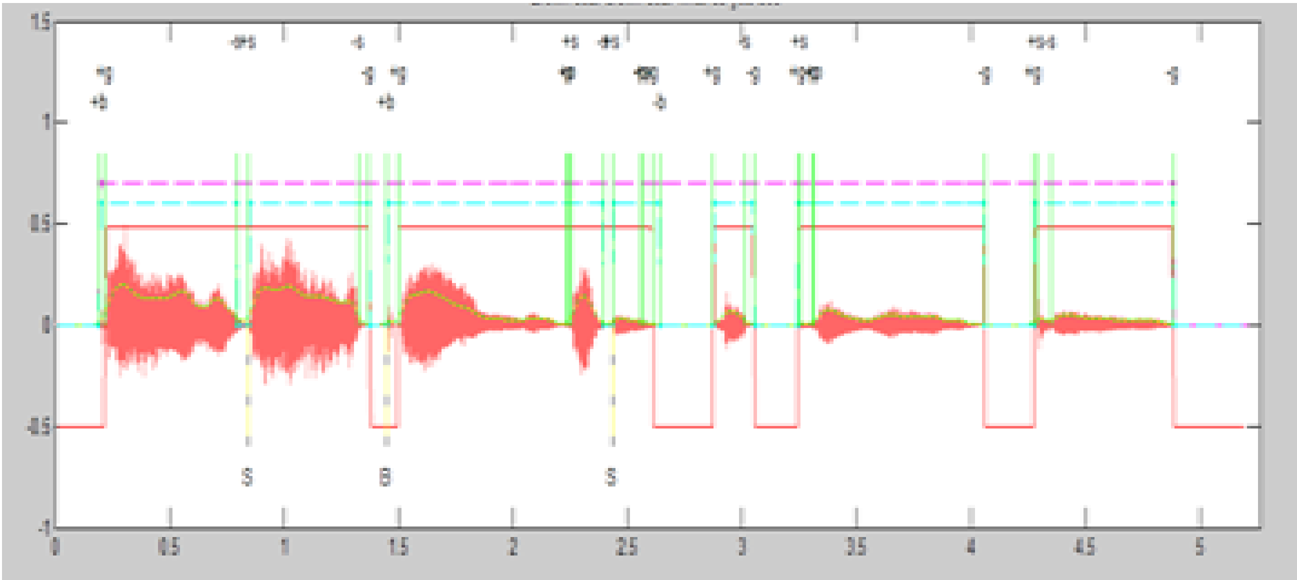
UNIVERSITY OF Cincinnati

Syllabic Clusters

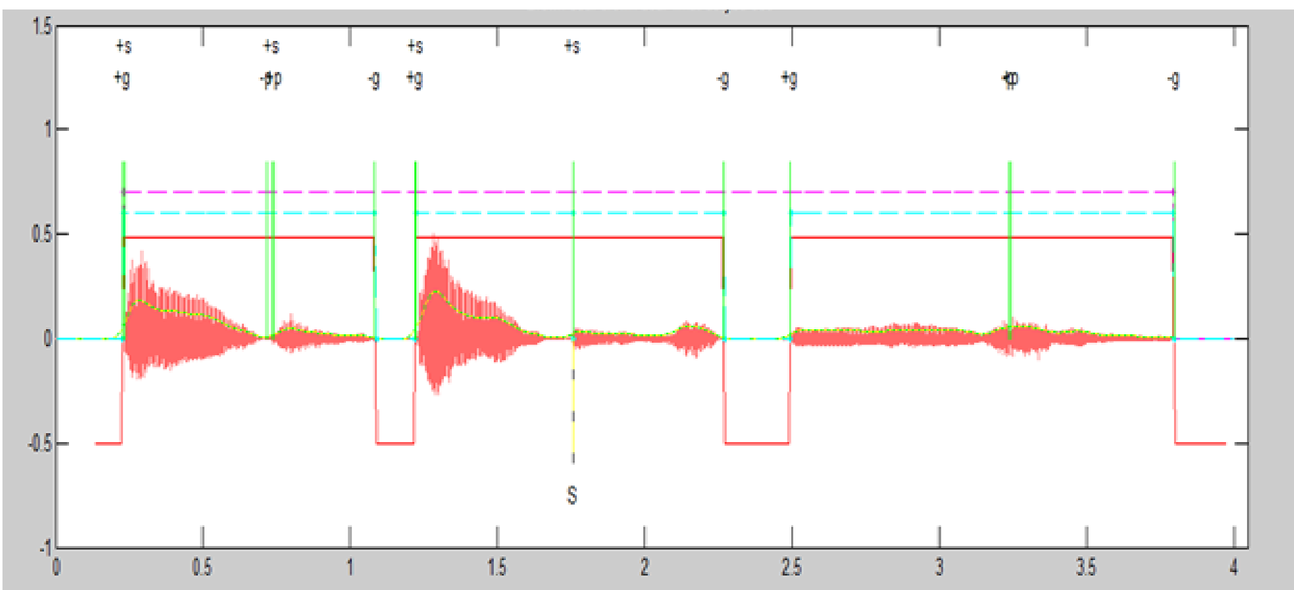
- The sequence and grouping of landmarks is related to how the speech was spoken.
- If spoken more canonically, as a string of intended syllables (dictionary form), more landmarks will be detected.
- If uttered less canonically, fewer landmarks will be detected.
  - less extreme movements
  - less precise timing
  - Reduced aerodynamic support

- V units correspond with a +g, -g sequence
- CV units such as “see” when precisely articulated may show up as +b, +g, -g
- CVC unit such as “bear” may appear as +b, -b, +g, -s,-g

UNIVERSITY OF Cincinnati



One utterance of normal child speech



One utterance of disordered child speech

Prior Studies

- Landmark and syllabic cluster analysis have been used to study:
  - Changes in infant babble across time to distinguishing infants who may be at risk for later communication disorders (Fell et al., 2002).
  - Normal vs. sleep deprived conditions (Boyce et al., 2008).
  - Parkinson's disease patients before and after undergoing Deep Brain Stimulation treatment (Chenausky, MacAuslan & Goldhor, 2011).
  - Clear vs. Conversational Speech (Boyce et al., 2013).
- Significant differences found in the number of landmarks detected and syllabic clusters as a result in change in age or condition (Fell et al., 1999, 2002; Boyce et al. 2011, 2013).

UNIVERSITY OF Cincinnati

Research Questions

RQ 1: Does the Landmark per Syllabic Cluster parameter predict speaker group (typical vs. disordered)?

RQ 2: Does the Syllabic Clusters per Utterance parameter predict speaker group (typical vs. disordered)?

RQ3: Does the Syllabic Cluster per Utterance parameter correlate with a conventional hand measure of syllabic complexity?

UNIVERSITY OF Cincinnati

Method

Participants

13 children (6 typical, 7 disorder status) age 3-6

Materials

Clinical Assessment of Articulation and Phonology 2nd Edition (Secord, Donohue, & Super Duper Publications, 2002)

- 46 single-words [Monosyllabic (bed) to multisyllabic (basketball)]
- 33 sentences elicited from the reading of children's book with repetitive language.
  - Ex. Brown bear, brown bear what do you see.

Number of words/sentences includes:

Stimuli	Typical (n= 6)	Disordered (n= 7)
Words	276	322
Sentences	266	299

UNIVERSITY OF Cincinnati

Method

Recording Conditions

- Speech samples were obtained in a quiet room at UC or at a site convenient to the participant.

- A Shure wireless microphone system used with an omnidirectional, subminiature, lavalier, condenser microphone.

- Samples are digitally processed at a sampling rate of 22K.

UNIVERSITY OF Cincinnati

Method

- Each token was phonetically transcribed and scored using the WORD COMPLEXITY MEASURE (WCM) (Stoel-Gammon,2010)

- WCM is scored across eight parameters in terms of word patterns, syllable structures, and sound classes to measure the complexity of each word.

Word Patterns

<two syllables

Stress on any syllable but the first

Syllable structures

Final consonant

Consonant Cluster

Sound Classes

Velar
Liquid
Rhotic
Fricative
Affricate

Method

Recordings were analyzed by the Syllabic Cluster algorithm from the SpeechMark Landmark Analysis System® MatLab tool box.

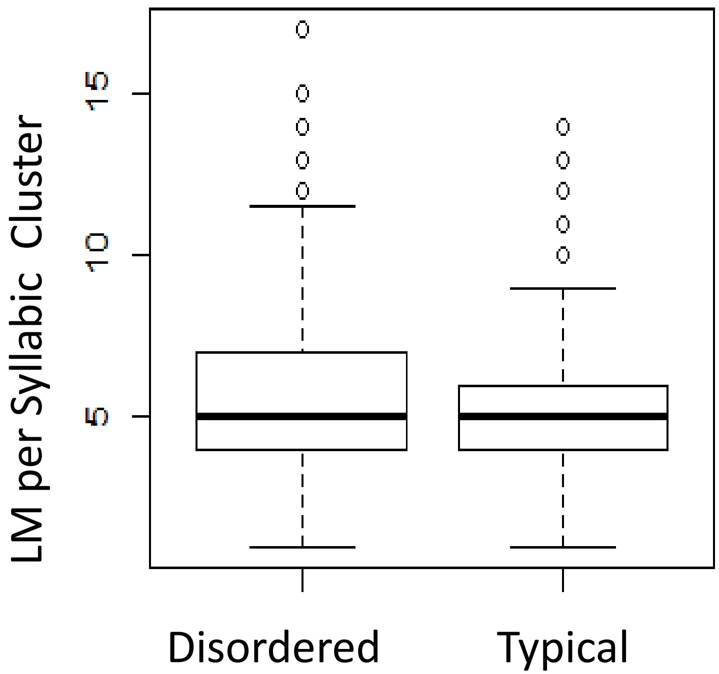
- Measures extracted were:
  - Total number of landmarks (LM)
  - Landmarks per Syllabic Cluster (Sylls)
  - Number of Utterances (Utts)
  - LMs per Sylls
  - Sylls per Utts.

UNIVERSITY OF Cincinnati

Results: RQ1

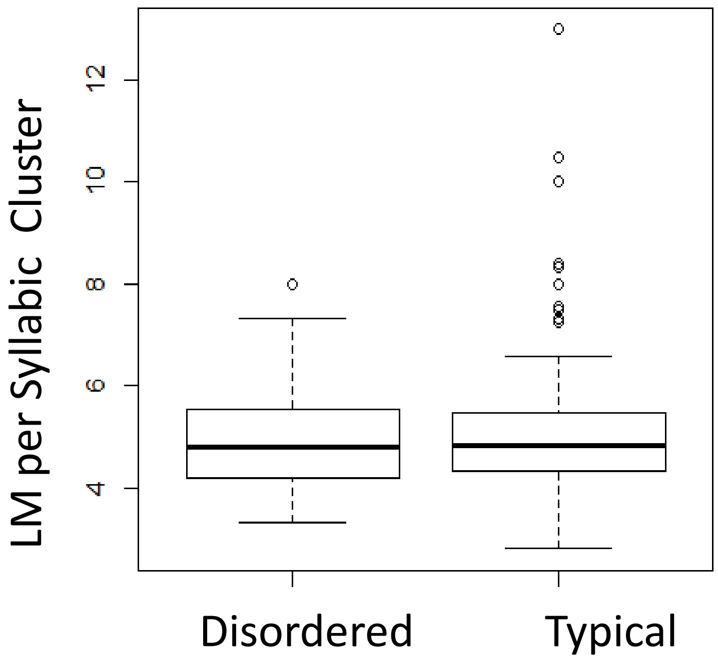
Does the Landmark per Syllabic Cluster parameter predict speaker group (typical vs. disordered)?

Words



p = 0.628

Sentences

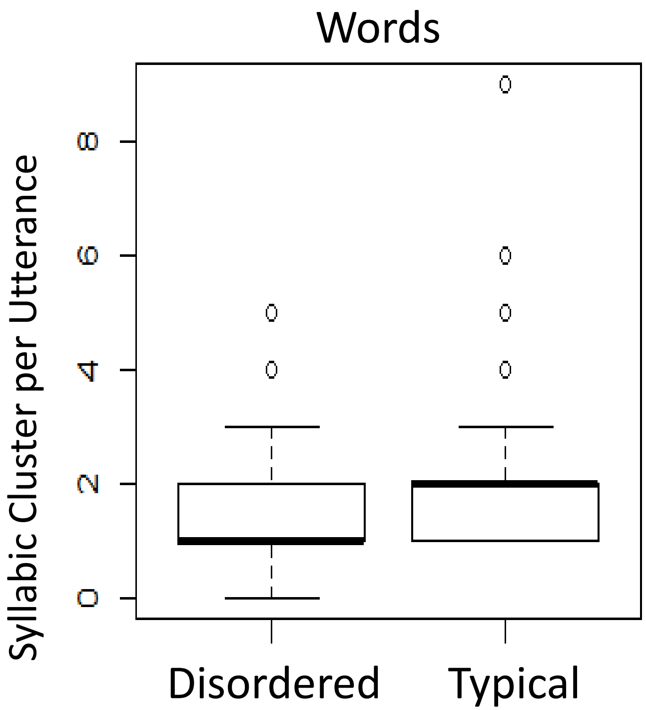


p = 0.1557

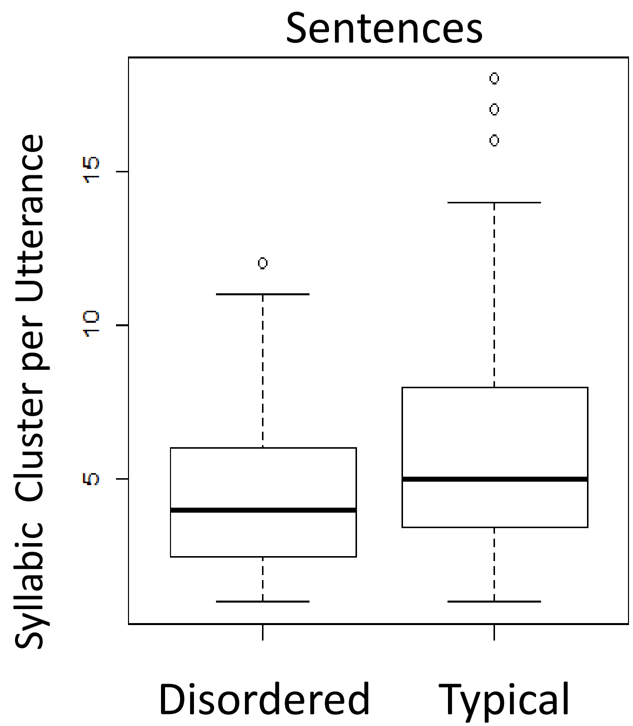
LM per Syll. does not significantly predict disordered status.

Results: RQ2

Does the Syllabic Clusters per Utterance parameter predict speaker group (typical vs. disordered)?



p=0.120

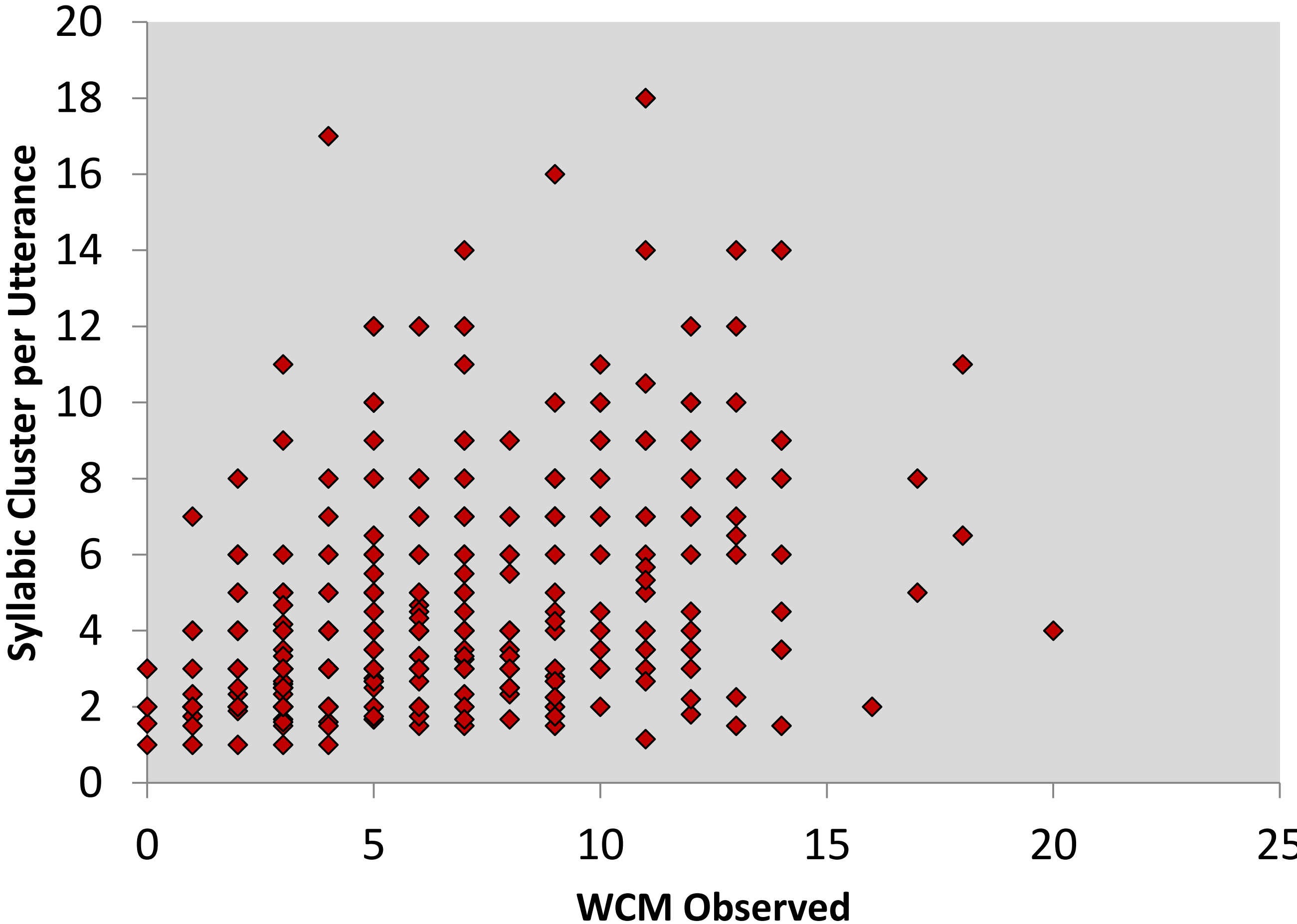


p = 0.0011

Syllabic cluster per utterance was a significant predictor of the disorder status in sentences.

RQ3

- Does the syllabic cluster per utterance parameter correlate with a conventional hand measure of syllabic complexity?



Moderate positive correlation between syllabic cluster per utterance and WCM observed in sentences, statistically significant, $r_s = 0.349$, $p < 0.001$.

Conclusions

- Syllabic cluster per utterance was a significant predictor of disorder status in running speech but not in words. This may be because the WORD sample included few multisyllabic words.
- Single word measures provide information on phonemic inventory but are limited in describing articulatory complexity intrinsic to running speech.
- Automated Syllabic Cluster per Utterance correlates with hand measures of word patterns and syllable structures.
- **Automated Syllabic Cluster detection is useful for measuring complexity of running speech samples without the need of phonetic transcription.**

UNIVERSITY OF Cincinnati

Proposed Future Direction

References

- Boyce, S., Fell, H. Wilde, L., MacAuslan, J. (2011): *Automated Tools for Identifying Syllabic Landmark Clusters that Reflect Changes in Articulation*", Proceedings from MAVEBA: 7th Annual Workshop for Models and Analysis of Vocal Emissions for Biomedical Applications
- Boyce S, Fell HJ, MacAuslan J. SpeechMark: Landmark Detection Tool for Speech Analysis. Paper presented at: INTERSPEECH2012. Nathani, S., Ertmer, D. J., & Stark, R. E. (2006). Assessing Vocal Development in Infants and Toddlers. *Clinical Linguistics & Phonetics*, *20*(5), 351–369.
- Fell, H. J., MacAuslan, J., Ferrier, L. J., & Chenausky, K. (1999). Automatic babble recognition for early detection of speech related disorders. *Behaviour & Information Technology*, *18*(1), 56-63.
- Fell, H., MacAuslan, J., Ferrier, L. J., & Worst, S., Chenausky, K., (2002). *Vocalization Age as a Clinical Tool*. Proceeding from ICSLP: International Conference on Speech Processing.
- Howitt, A. W., *Automatic Syllable Detection for Vowel Landmarks*, doctoral thesis M.I.T., Cambridge,MA. 2000.

UNIVERSITY OF Cincinnati

References

- Liu S. A. (1994)Landmark detection for distinctive feature-based speech recognition. *The Journal of the Acoustical Society of America.*;100(5):3417-3430.
- Oller, D. K. (2000). *The emergence of the speech capacity*. Psychology Press.
- Stoel-Gammon, C. (2010). The word complexity measure: Description and application to developmental phonology and disorders. *Clinical linguistics & phonetics*, *24*(4-5), 271-28
- Stevens, K.N., Manuel, S., Shattuck-Hufnagel and Liu, S. 1992. Implementation of a model for lexical access based on features. *Proc. Int'l. Conf. Spoken Language Processing*, Banff, Alberta, **1**, 499-502.
- Stevens, K. N. (2000). Diverse acoustic cues at consonantal landmarks . *Phonetica, 57*, 139-151.
- Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustic Society of America, 111*, 1872-1891.

UNIVERSITY OF Cincinnati