A New & Improved Drift4 for Performative Speech Analysis

Sarah Yuniar   **November 14, 2022**

SPOKENWEBLOG | audio, data collection, data visualization, Drift, drift4, Gentle, literary audio, marit MacArthur, prosody, software, transcription, voxit, words per minute, WPM

Tweet   Share   Cite this article   "

Introduction

Drift is a highly accurate pitch-tracker prototyped in 2016 by Robert Ochshorn and Max Hawkins. Its further development has been supported by a NEH Digital Humanities Advancement grant and now by SpokenWeb. At UC Davis, undergraduate research assistants Sarah Yuniar and Hannan Waliullah, working with Marit MacArthur and Lee M. Miller, have beautifully improved its functionality and interface.

Drift measures what human listeners perceive as vocal pitch (the fundamental frequency, the vibration of the vocal cords, as measured in hertz) every 10 milliseconds in a given recording, visualizing it in an easy-to-read, horizontally scrolling pitch trace, aligned with the text being read. Drift uses an algorithm developed by Byung Suk Lee and Daniel P. W. Ellis at Columbia University to work with precise accuracy on the noisy, low-quality vocal recordings common in the audio archive. Additionally, Drift incorporates the forced alignment features of Gentle, developed by Robert Ochshorn and Max Hawkins, which aligns a given transcript with an audio file's pitch trace.

New User Interface

Drift4 is updated with a new user interface designed with ease of use in mind, allowing anyone from experts to laypeople to view pitch, timing and intensity data about a given recording in an informative yet intuitive way. It also includes extensive instructions and tutorials. Prosodic measures include hoverable previews of their definition, with additional descriptions found on the site's "About" page.

Drift4's new user interface, using a reading of The Woman Syndrome by Dorothy Livesay.

Voxit prosodic measures have easily accessible descriptions when hovered over.
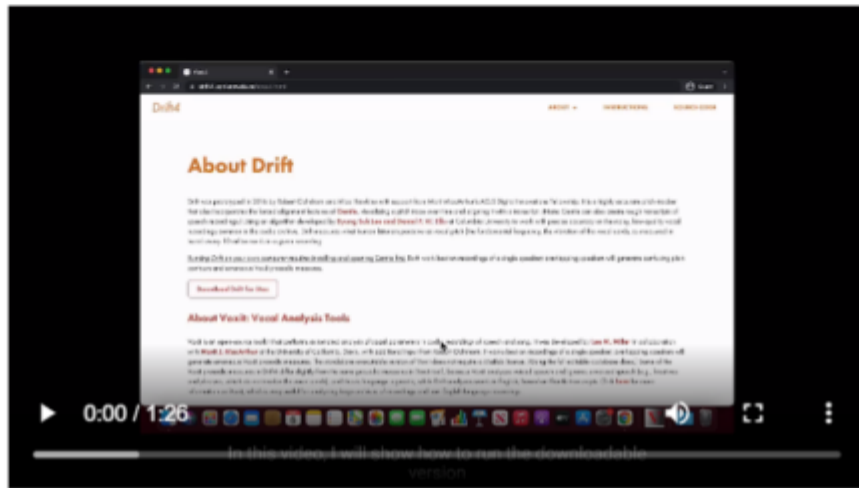
Addition of Instructions

Instructions are provided on a separate sub-page to help new users navigate Drift's wide range of features, many of which were added only recently in Drift4, like draggable document lists, additional downloadable data, and auto-scrolling.

Instructions

Video Tutorials

Launching the Drift application on Mac [Mac]

Uploading files to Drift [Mac] [Windows] [Web]
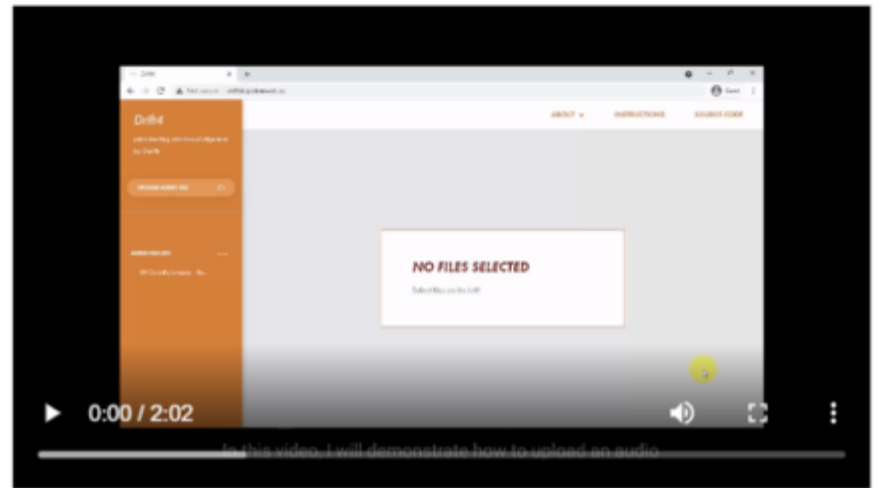
Instruction videos are provided on the web and app versions of Drift4.

Voxit

Drift now incorporates most of the same prosodic measures as Voxit, a distant listening toolkit for calculating prosodic measures, developed by Lee M. Miller in collaboration with Marit J. MacArthur and with additional input from Robert Ochshorn. These measures include *Drift f0 Mean Absolute Velocity, Gentle Complexity All Pauses*, and more. (WPM is included in Drift but not in Voxit, as Voxit measures voiced periods, not words.) While Drift is a slow listening, qualitative tool for looking at a few recordings that provides quantitative data, Voxit is more useful for distant listening, as it can process a large number of recordings. The incorporation of Voxit prosodic measures in this newest version of Drift allows portability to those who have used Voxit in the past, and provides a more standardized and consistent approach to prosodic measures. These measures can be calculated over selected time durations, allowing the user to study how an audio recording—and a speaker's vocal performances—changes over time.

	WPM	Drift f0 Mean (hz)	Drift f0 Range (octaves)	Drift f0 Mean Abs Velocity (octaves/sec)	Drift f0 Mean Abs Accel (octaves/sec^2)	Drift f0 Entropy	Dynamism	Intensity Mean Abs Velocity (decibels/sec)	Intensity Mean Abs Accel (decibels/sec^2)	Intensity Segment Range 95 Percent (decibels)	Gentle Mean Pause Duration (sec)	Gentle Pause Rate (pause/sec)
full recording duration	119	271.71	2.75	1.38	0.74	3.55	9.22	194.85	191.06	21.67	0.65	0.45
selection	135	279.21	2.75	1.34	0.74	3.48	9.1	185.8	181.89	23.76	0.66	0.35

A snapshot of some of the prosodic measures calculated using Voxit.

Windowed Data

Alternatively, users can download a CSV representing prosodic measures within windows of time. This windowed representation simply evaluates the same measures over every twenty-second segment of the audio recording. The twenty-second length is the result of rigorous testing, which showed that it provides measures that are both stable and short enough to track typical prosodic style changes over time (shorter window lengths caused the values to vary too much). Thus, the values for 20-second windows are fairly reliable in characterizing patterns, tendencies, and dynamics in a given recording.

	A	B	C	D	E	
1	measure_label	window_len	full_transcript	seg_1	seg_2	
2	WPM	20	119	126	95	
3	Drift_f0_Mean_(hz)	20	271.7071925	265.636735	266.292485	
4	Drift_f0_Range_(octaves)	20	2.749999002	2.749999	1.20831897	
5	Drift_f0_Mean_Abs_Velocity_(octaves/sec	20	1.38390541	1.28102538	1.47703525	
6	Drift_f0_Mean_Abs_Accel_(octaves/sec^2)	20	0.739007382	0.71937732	0.74086737	
7	Drift_f0_Entropy	20	3.545917821	3.4209523	3.6340748	
8	Gentle_Mean_Pause_Duration_(sec)	20	0.65	0.68	0.59	
9	Gentle_Pause_Rate_(pause/sec)	20	0.448	0.3	0.529	
10	Gentle_Complexity_All_Pauses	20	6.981676231	6.44727604	10.2622104	
11	Gentle_Pause_Count_>100ms	20	16	6	10	
12	Gentle_Pause_Count_>500ms	20	8	3	4	
13	Gentle_Pause_Count_>1000ms	20	3	1	2	
14	Gentle_Pause_Count_>2000ms	20	0	0	0	
15	Gentle_Long_Pause_Count_>3000ms	20	0	0	0	

An example of windowed data, featuring prosodic measures of two 20-second long segments for *The Woman Syndrome*.

In combination, Drift, Gentle, and Voxit work together to assist the study of audio recordings in a more intuitive way than ever before. Questions like how dynamic a speaker's voice is can be inferred from Drift's pitch trace visualization, but can be quantified using Voxit's "dynamism" measurement, along with "pitch velocity/acceleration". Similarly, the speed of the speaker's vocal pitch changes can reveal whether or not they are nervous or excited, and can be visually interpreted through the alignment of the transcript on the pitch trace or concluded quantitatively using the "WPM" measurement.

Drift can be developed further, with potential for additional prosodic measures like emotional data, according to the needs and interests of SpokenWeb. Our hope is that SpokenWeb members will begin using Drift more, in both research and teaching, and will provide feedback on other features we might develop. Drift is a wonderful tool to realize SpokenWeb's mission of bringing interpretability to oral literature. For more explanation of how Drift and Voxit can be used in analyzing literary recordings, readers may be interested in these publications using the tools. If you know of other publications using the tools, please share them and we will add them. Please contact Marit MacArthur with any questions and suggestions at mjmacarthur@ucdavis.edu.

References

- MacArthur, Marit, Rambsy, Howard, II, Wu, Xiaoliu, Ding, Qin, and Miller, Lee M. "101 Black Women Poets in Mainly White and Mainly Black Rooms." *Los Angeles Review of Books*, August 27, 2022.
- "John Ashbery's Reading Voice." Oct. 29, 2019. *The Paris Review Online*.
- MacArthur, Marit and Miller, Lee M. "After Scansion: Visualizing, Deforming and Listening to Poetic Prosody." *Stanford ARCADE Colloquy Series: Alternative Histories of Prosody*, Dec. 13, 2018. [Essay and podcast]
- MacArthur, Marit, Zellou, Georgia and Miller, Lee M. "Beyond Poet Voice: Sampling the Performance Styles of 100 American Poets." *Journal of Cultural Analytics*, March 2018.

Sarah Yuniar

Sarah Yuniar is a student at the University of California, Davis, pursuing a B.S. in Computer Science and Engineering. She is currently an undergraduate assistant to Dr. Marit J. MacArthur and Dr. Lee Miller, contributing to the development of Drift4.

About Us

Contact Us

Get Involved

Downloads

News

Opportunities

Submit an Event

Admin

SpokenWeb is a SSHRC-funded partnership grant.

All material that appears on this website is used for the purposes of academic research and critical study.

SSHRC CRSH

© 2010-2023. All rights reserved.