Description of Inter-Item Timing

Tyler Peckenpaugh

3/25/2019

Contents

Inter-item timing

Inter-reading time (IRT) is a measure of the amount of time inbetween when a subject stops speaking during a cold reading and when they begin speaking for a previewed reading. Practically, this was done over 1,533 recording (33 participants, 48 items = 1584 pairs, with some missing data resulting in the 1,533). This was measured using Google's WebRTC Voice Activity Detection (VAD).

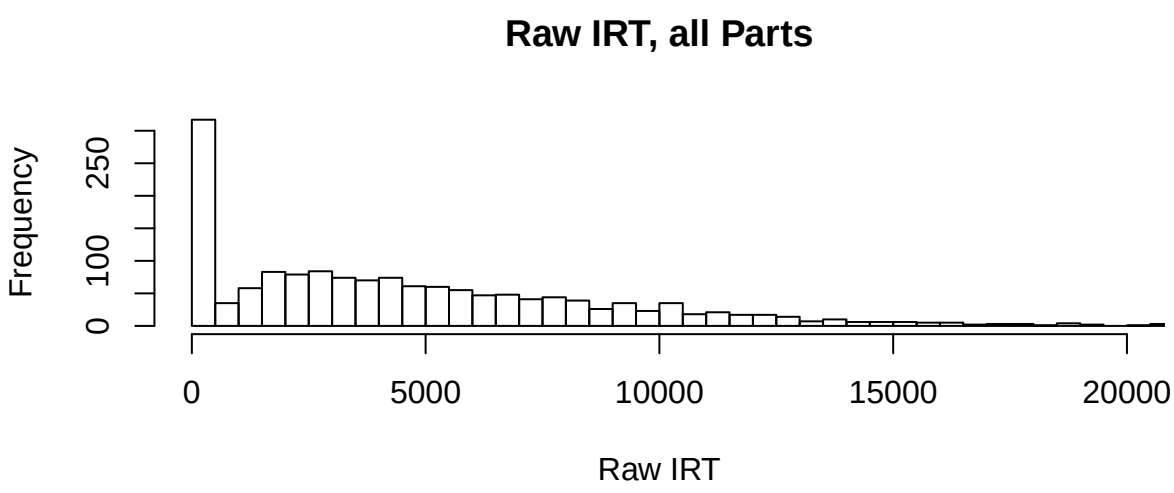
> IRT(item) = the timestamp of the start of speech for item[reading one] plus (the length of item[reading two] minus the timestamp of the end of speech)

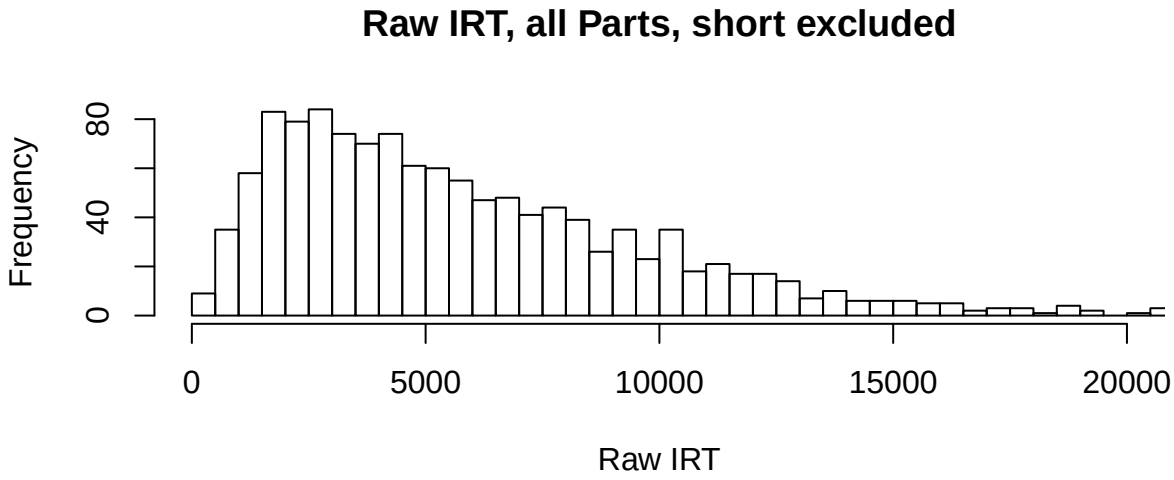
Subjects were asked to read each sentence twice, once with no preview at all, and then again after unlimited preview. IRT is meant to represent their chosen preview time for the second reading.

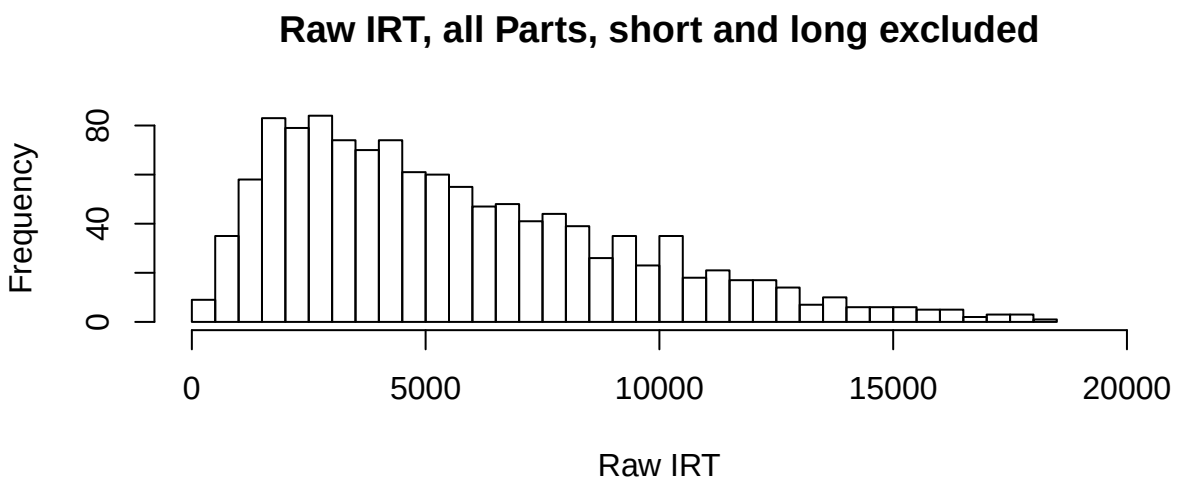
Distribution of IRTs, all participants

The overal mean IRT for all participants, all items (including fillers), and all conditions is 4765.5674629ms (4473.9352857).

The following histograms show the distribution of IRT across all items and all participants. In the second graph, overly short IRTs (shorter than 150ms) are excluded. In the third, overly long (longer than ~18s) and overly short IRTs are excluded.


Raw IRT, all Parts
Frequency
0
100
250
0
5000
10000
15000
20000
Raw IRT


Raw IRT, all Parts, short excluded
Frequency
0
40
80
0
5000
10000
15000
20000
Raw IRT



Missing data and attrition

Due to noise in recordings and/or technical difficulties during data collection, a number of IRTs are missing in the data. The following total shows which participants are missing how many IRTs; ideally each would have 48 IRTs.

Participants with fewer than 72% of the expected number of IRTs are excluded: 7, 13, 5, 17, 12, 21, 22, and 3, 8 in total. The resulting number of remaining participants is 25.

Group sizes after attrition

The following table shows how the participants are distributed across groups after attrition. While each group and each split are similar in size overall, there is an unfortunate disparity across ordering for two groups: Group 2 and Group 4.

Avaialability of experimental item IRT after attrition

For the remaining participants, the availability of IRT for experimental items was 88.7%, ranging by participant from 16/16 to 10/16.

In detail, 13 of 25 participants had all experimental IRTs available, while five had 11/16. There were two participants who had 12/16, 14/16 and 15/16 respectively, and one with 10/16.

Distribution of experimental item IRT after attrition

The following histograms show the distribution of experimental item IRTs after attrition, and then the Winsorized IRTs, and finally the common log of winsorized IRTs, which are the shape of the data most suited to regression analyses.

Table 1: Missing IRTs

Participant	IRTs missing	IRTs availble	Recording pairs	Percentage available	Group
7	45	3	48	6.25%	EXCL
13	43	5	48	10.42%	EXCL
5	7	18	25	37.50%	EXCL
17	29	19	48	39.58%	EXCL
12	27	21	48	43.75%	EXCL
21	27	21	48	43.75%	EXCL
22	24	24	48	50.00%	EXCL
3	23	25	48	52.08%	EXCL
2	13	35	48	72.92%	1-2
4	13	35	48	72.92%	1-0
9	13	35	48	72.92%	1-1
11	1	35	36	72.92%	2-3
14	10	38	48	79.17%	2-2
6	8	40	48	83.33%	1-2
203	8	40	48	83.33%	1-3
8	6	42	48	87.50%	1-0
16	5	43	48	89.58%	2-0
15	2	46	48	95.83%	2-3
19	2	46	48	95.83%	2-3
204	1	47	48	97.92%	1-0
210	1	47	48	97.92%	2-2
1	0	48	48	100.00%	1-1
10	0	48	48	100.00%	2-2
20	0	48	48	100.00%	2-0
201	0	48	48	100.00%	1-1
206	0	48	48	100.00%	1-2
207	0	48	48	100.00%	2-3
208	0	48	48	100.00%	2-0
209	0	48	48	100.00%	2-1
210	0	48	48	100.00%	2-2
212	0	48	48	100.00%	2-0
214	0	48	48	100.00%	2-2
215	0	48	48	100.00%	2-3

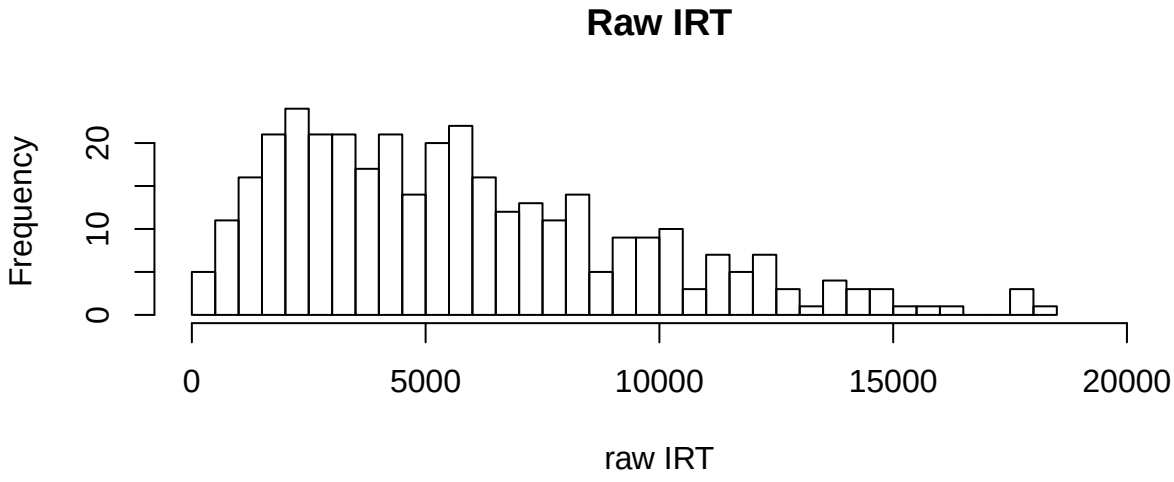
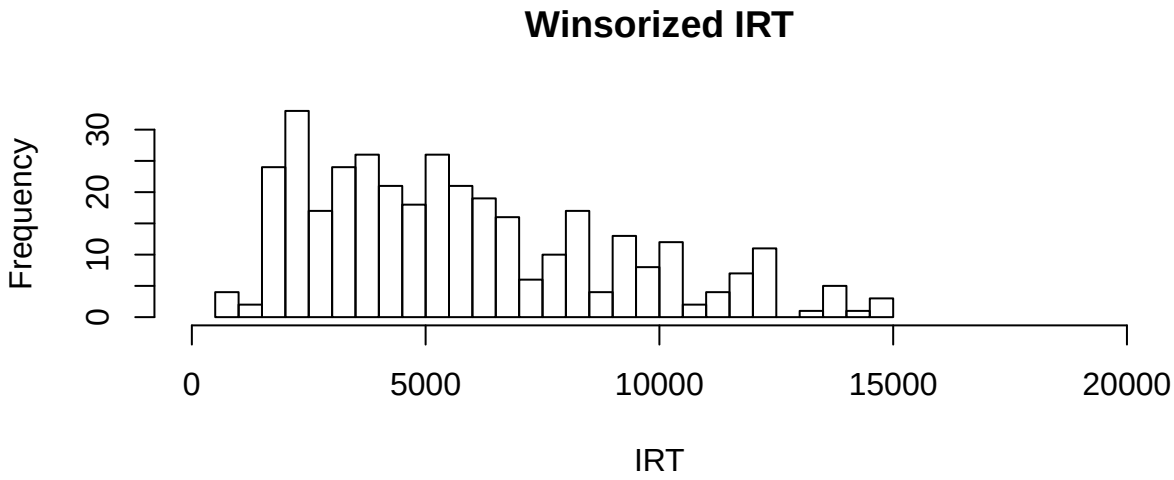
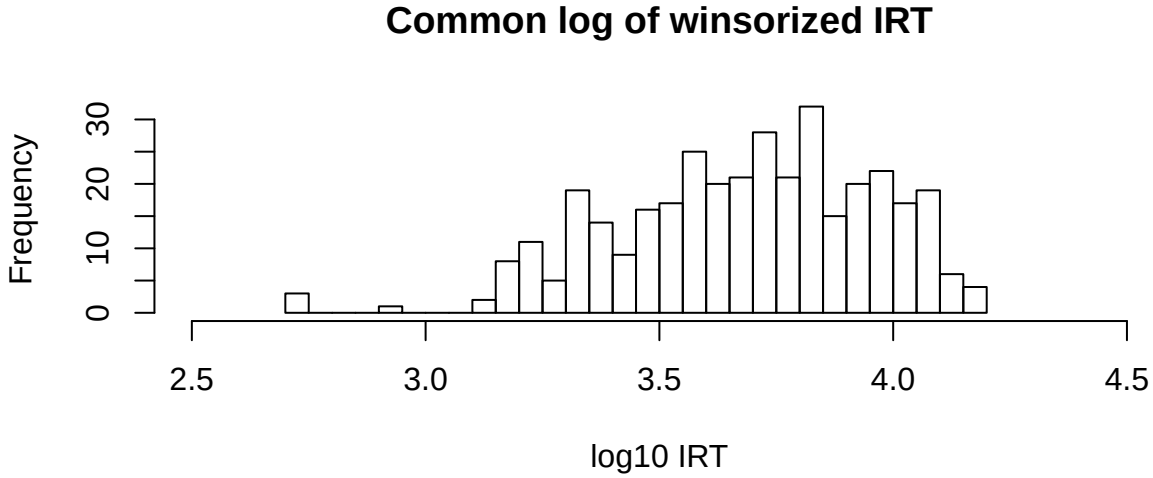


Table 2: Group/order totals after attrition

	Split AB	Split BA	Group Total
Group 1	3	4	7
Group 2	4	1	5
Group 3	3	4	7
Group 4	1	5	6
Split Total	11	14	25





Mean and SD of winsorized IRT by condition

If we assume that interrogative PP-attachment garden paths are easier to process as an interrogative than in the declarative, and that IRT represents how difficult a sentence is to process, we would expect the difference in mean IRT to be larger for declarative garden paths compared to declarative controls than for the same comparison of interrogatives.

Table 3: Condition means

	Mean	SD
-Q -GP	5291.473	3027.415
-Q +GP	5857.861	3106.019
+Q -GP	5799.589	3182.036
+Q +GP	6086.307	3699.116

Table 4: Mean IRT by condition and participant

participant	-Q -GP	-Q +GP	+Q -GP	+Q +GP	pattern
1	1666.500	2039.750	2057.000	1741.500	TRUE
2	6587.100	6958.250	5279.600	6744.000	TRUE
4	2125.000	4274.333	7605.667	7016.667	TRUE
6	13194.333	12307.333	12414.000	12747.333	FALSE
8	4297.267	4113.467	3254.700	5018.600	FALSE
9	1289.000	1531.000	1920.000	610.000	TRUE
10	9008.500	9855.500	9193.500	9599.750	TRUE
11	6418.000	7440.000	5546.333	5907.667	TRUE
14	8135.333	10948.000	8456.000	12277.750	TRUE
15	7489.050	7954.500	8565.067	7734.050	TRUE
16	7376.500	8180.950	10500.667	9452.533	TRUE
19	4125.450	8947.333	9201.667	13855.600	TRUE
20	10183.500	9135.000	11905.500	10141.500	FALSE
201	3874.750	5004.000	4674.000	4145.250	TRUE
203	1828.900	1785.667	1799.867	1966.400	FALSE
204	3829.500	4672.150	6241.150	4841.533	TRUE
205	7686.500	10064.500	7309.750	10237.500	TRUE
206	4754.000	4401.500	4394.000	4580.250	FALSE
207	5440.250	5959.500	6504.000	5119.500	TRUE
208	3006.500	3185.000	2614.000	2922.250	TRUE
209	3047.750	3468.750	3182.000	3140.000	TRUE
210	5810.500	4905.250	5946.000	5061.250	FALSE
212	3696.500	5033.250	4170.750	5323.750	TRUE
214	3542.750	2977.500	3201.000	2699.500	FALSE
215	4699.000	4284.000	4870.250	4501.500	FALSE

The means of the Winsorized IRT by condition indeed show this pattern.

Difference for declaratives: 566.39; for interrogatives: 286.72. While the difference across the interrogative/declarative condition is not huge, it is there.

Item and subject variation

There is variation across participants in terms of whether or not they show this pattern.

Number of participants who show predicted pattern

17 of 25 participants show the expected pattern (p = 0.05 in a one-tailedd binomial test). This suggests that the pattern generalizes across participants.

Table 5: Mean IRT by condition and item

item	-Q -GP	-Q +GP	+Q -GP	+Q +GP	pattern
1	4075.250	6537.250	6718.914	7744.833	TRUE
2	4108.680	5402.600	3838.200	6312.320	TRUE
3	5673.629	6396.167	3856.000	6992.867	TRUE
4	7449.667	5254.286	4991.767	4893.250	FALSE
5	4342.167	7074.500	3792.600	4460.200	TRUE
6	6443.667	5496.286	6402.833	4428.800	FALSE
7	4835.167	5710.000	6429.367	7214.233	TRUE
8	5640.267	4842.200	6494.333	5517.486	FALSE
9	4453.750	5218.600	6816.333	7199.333	TRUE
10	5689.400	8060.714	3827.400	5380.040	TRUE
11	4081.067	3816.800	7820.000	6583.600	FALSE
12	3372.800	5235.200	6204.714	8042.500	TRUE
13	5249.120	4808.750	5446.200	6104.400	FALSE
14	6262.700	5973.800	5414.650	4457.800	FALSE
15	3340.000	6908.833	5950.000	7240.571	TRUE
16	8115.000	5837.971	7702.400	2976.200	FALSE

Number of items that show predicted pattern

Notably, only 9 of 16 items show the pattern ($p = 0.4$ in a one-tailed binomial test). This suggests that the pattern does not generalize across items. This is likely problematic for any regression analyses.