CORRIGENDUM: This corrigendum contains a reanalysis of "Pouw, W., & Dixon, J. A. (2019). Quantifying gesture-speech synchrony. In the *6th gesture and speech in interaction conference* (pp. 75-80). Universitaetsbibliothek Paderborn.". The reanalysis is performed with corrected pitch time series data which was mistakenly thresholded by the first author at a range of 75-125Hz, in contrast to the reported 75-500Hz. Thus, we report main differences in results relative to the original report, and we conclude that the main interpretations remain unchanged (e.g., peak velocity is most closely aligned on average to peak in F0; synchronization between speech and gesture is apparent).

CORRIGENDUM: Quantifying gesture-speech synchrony

Wim Pouw [1, 2]
Donders Center for Brain, Cognition, and Behavior, Radboud University Nijmegen

Author note: Correspondence should be addressed to Wim Pouw (wim.pouw@donders.ru.nl).
CORRIGENDUM Open data: https://osf.io/5ja6y/
Acknowledgements: We would like to thank Rüdiger Fasching for finding the discrepancy of the F0 bandwidths settings in our data.

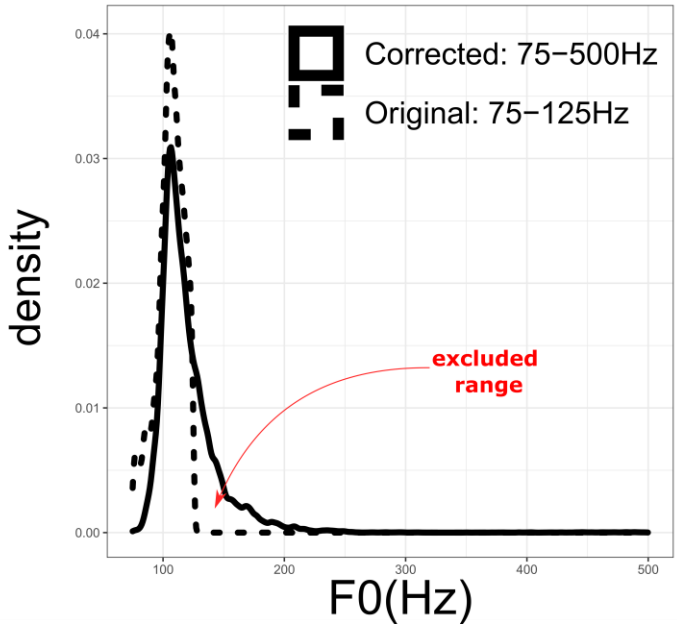
Corrigendum

In the original report (Pouw & Dixon, 2019) the extraction of the Fundamental Frequency (F0) was reported to be performed with a bandwidth of 75-500Hz. However, upon inspection of the time series with a re-extraction of F0 using PRAAT with 75-500Hz bandwidth, it was observed that the original acoustic data were accidentally thresholded at 75-125Hz for F0 extraction. This may change our analysis and interpretation as the sensitivities to pitch tracking are different, which may lead to differences in the temporal approximation of local F0 maxima. Therefore, we have corrected the pitch time series (we have provide these corrected files on the Open Science Framework), and we report here a comparison with the original findings and the corrigendum reanalysis. We conclude that there are numerical changes, but we see no major changes to the interpretation of our results.

Main effect of corrected bandwidth settings on F0 estimations

In the reported experiment all participants were males, which have an average F0 range of about 80-180Hz. Thus, it is likely that by setting our bandwidth to an upper bound of 125Hz that these vocalizations are not approximated appropriately. In Figure C1 we can indeed see that higher F0 values are observed when the bandwidth is corrected to the 75-500Hz. This warrants a reanalysis of the data, as our peak analysis is based on local maxima of the F0 time series.

Figure C1. F0 distributions for the corrected and the original dataset



Note. Density distribution showing the difference in observed F0 values for all the participants depending on PRAAT bandwidth settings. It can be seen that there is a sudden drop-off of the occurrence of F0 values higher than 125Hz; this is due to the incorrect bandwidth settings in the original study.

Reanalysis

The main numerical results of the timing differences between kinematic parameters in peaks in F0 are reported in TABLE B. Below is a comparison with the corrected results. Our main conclusion that peak velocity is on average the most closely aligned with peak F0 remains unchanged (See also Figure C2 with comparisons of the original distributions, versus the corrected distributions). There are however a changed results concerning the variability in gesture-speech timing per gesture type. In contrast to the clear lack of differences between timings for gesture timing, we do now find reliably F0 peak timing differences with gesture velocity, and deceleration peaks.

CORRIGENDUM Table B. Mean Difference, *D*, in milliseconds (peak pitch – gesture property)

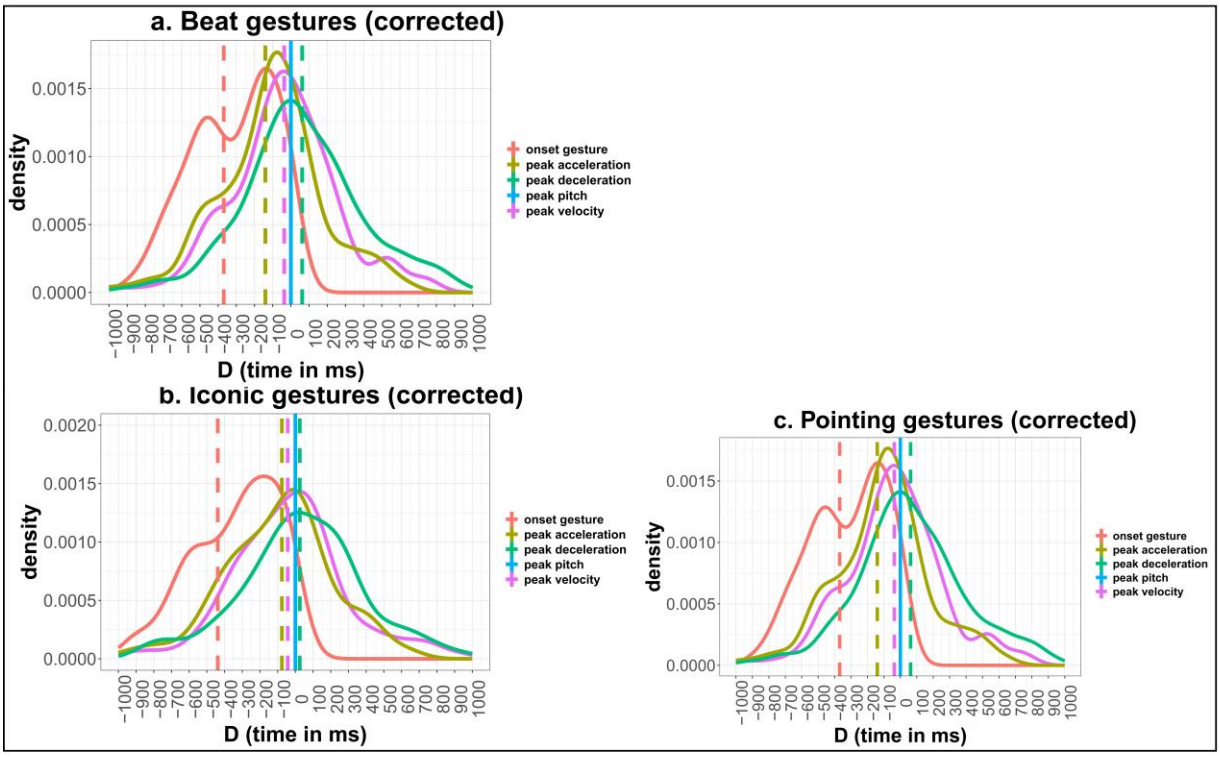
Kinematic propert	BEAT *M* (SD) 95% CI [lower, upper]	ICONIC *M* (SD) 95% CI [lower, upper]	Pointing *M* (SD) 95% CI [lower, upper]	F-test Diff *p* (corrected; *p**5) Bayes Factor
Onset *M* (SD) 95% CI [lower, upper]	-349 *(314)* [-400, -299]	-449*(426)* [-526, -374]	-379 (427) [-536, 223]	*F* (1, 3) = 1.28 *p* > .99 BF_{01} = 3.04
CORRIGENDUM Onset	-365 *(325)* [-417, -312]	-434 *(524)* [-528, -341]	-369 (326) [-422, 317]	*F* (1, 3) = 7.44 *p* = .361 BF_{01} = 0.71
Peak velocity	-22 *(334)* [-75, 31]	-55 *(589)* [-159, 49]	-59 (349) [-185, 67]	*F* (1, 3) = 1.39 *p* > .99 BF_{01} = 3.08
CORRIGENDUM Peak velocity	-36 *(325)* [-87, 16]	-41 *(494)* [-129, 45]	-36 (325) [-88, 15]	*F* (1, 3) = 49.87 *p* = 0.029 BF_{01} = 0.284
Peak acceleration	-126 *(377)* [-187, -66]	-87 *(630)* [-199, 23]	-144 (380) [-282, -7]	*F* (1, 3) = 0.612 *p* > .99 BF_{01} = 7.69
CORRIGENDUM Peak acceleration	-139 *(353)* [-196, -83]	-74 *(508)* [-164, 15]	-140 (354) [-196, -84]	*F* (1, 3) = 5.75 *p* = .48 BF_{01} = 2.151
Peak deceleration	81 *(367)* [22, 138]	10 *(499)* [-78, 98]	4 (344) [-120, 129]	*F* (1, 3) = .718 *p* = .104 BF_{01} = 0.22
CORRIGENDUM Peak deceleration	67 *(354)* [11, 124]	23 *(501)* [-65, 112]	62 (357) [6, 119]	*F* (1, 3) = 63.67 *p* = .021 BF_{01} = 0.06

Note. P-values are Bonferroni adjusted (for 5 repeated comparisons) for four comparisons. We have also computed Bayes Factors (BF) with non-informative default prior widths p(M) = 0.5 using R package "BayesFactor" (Rouder, Morey, Speckman, & Province, 2002) as to provide a measure of evidence for the null-hypothesis over the alternative hypothesis. The BF's are reported for the likelihood of the observed data provided the null-hypothesis true (no differences per gesture type) versus the alternative hypothesis (BF_{01}). If BF_{01} > 3 this can be treated as moderate/substantial evidence (3-10 as

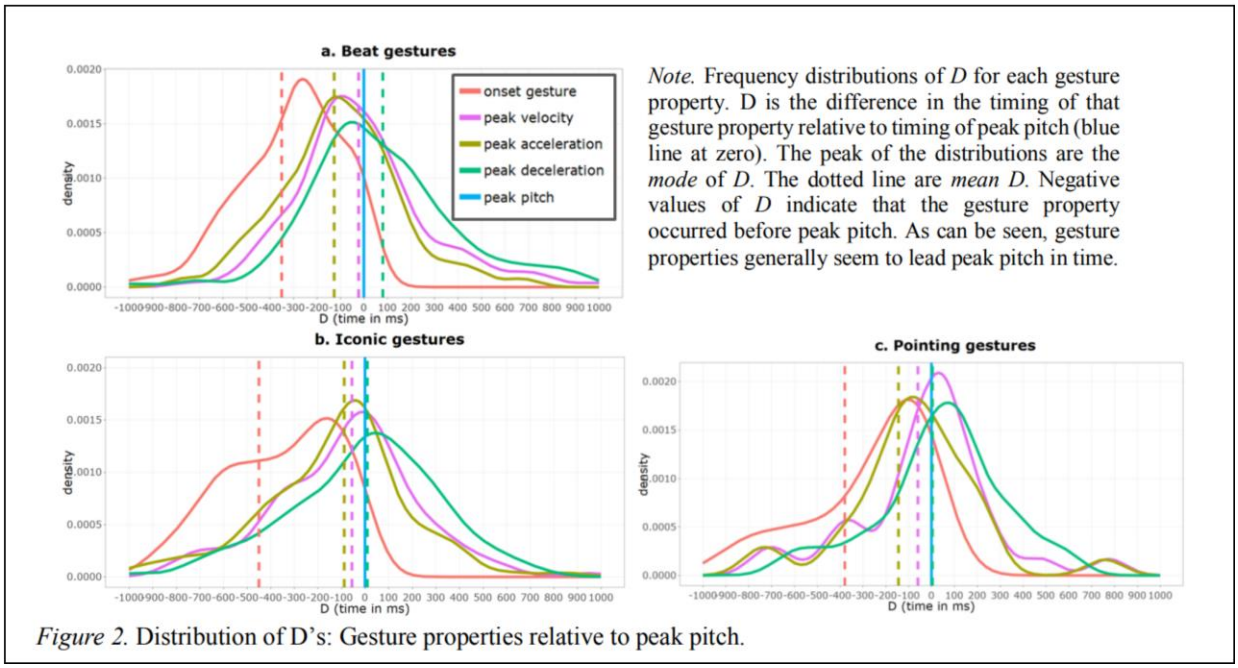
strong evidence; see Rouder, Morey, Verhagen, Swagman, Wagenmakers, 2016), and can be read as the observed data being 3 (or more) times more likely under the null-model versus the alternative model.

Figure C2. Comparison main timing figures (corrected versus original)

corrected



Original



Note. Frequency distributions of *D* for each gesture property. D is the difference in the timing of that gesture property relative to timing of peak pitch (blue line at zero). The peak of the distributions are the *mode* of *D*. The dotted line are *mean D*. Negative values of *D* indicate that the gesture property occurred before peak pitch. As can be seen, gesture properties generally seem to lead peak pitch in time.

Figure 2. Distribution of D's: Gesture properties relative to peak pitch.

Main Statistical Test

Below in red we provide the corrected values for the statistical tests we performed to assess differences in gesture kinematic peaks (gesture properties) relative to F0. As can be seen, the differences in parameter values do not produce a difference in interpretation.

We found that these gesture properties differed reliably in their D's, F (2, 6) = 17.54, p < .005 (D's, F (2, 6) = 18.13, p < .005). Paired post-hoc comparisons (p-values Bonferroni corrected) revealed that peak velocity shortly led peak pitch (MD = -39, SDD = 454, 95%CI[-90: 11]) (MD = -57, SDD = 385, 95%CI[-93: 21]), as compared to peak deceleration which followed peak pitch (p < .001; MD = 44 , SDD = 424, 95%CI[-3 : 92]) (p < .001; MD = 35, SDD = 385, 95%CI[-2: 74]). Peak acceleration was furthest from peak pitch (MD = -113, SDD = 494, 95%CI[-168 : -58]) (MD = -152, SDD = 373, 95%CI[-189: 115]), and was statistically different from peak velocity and peak deceleration (ps < .001) (p < .001).

Conclusion

In our original report we performed peak analysis, where we determine local maxima based on the values of the timeseries of F0 and some kinematic parameter. Because we have mistakenly set the upper bound of the F0 tracking by Praat at 125Hz as opposed to 500Hz, it was possible that the actual F0 ranges of the male speakers in our sample was not correctly tracked. Therefore, we have performed a reanalysis with the corrected pitch track. We concluded in our main report that gesture-speech synchrony was apparent given the clear peaked timing distributions; that peak velocity on average tends to align with F0, as compared to other kinematic parameters; we show how kinematic parameters relate to peak F0 in terms of their sequence. These key interpretations remain unchanged.

References

Pouw, W., & Dixon, J. A. (2019). Quantifying gesture-speech synchrony. *Proceedings of the 6th Meeting of Gesture and Speech in Interaction*, 68–74. https://doi.org/10.17619/UNIPB/1-812