Journal of Experimental Psychology: Learning, Memory, and Cognition

More Laws for Pauses: Replication and Generalization

Taylor M. Mezaraups and David L. Gilden

CITATION

Mezaraups, T. M., & Gilden, D. L. (2023, August 10). More Laws for Pauses: Replication and Generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://dx.doi.org/10.1037/xlm0001280

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Learning, Memory, and Cognition

ISSN: 0278-7393

https://doi.org/10.1037/xlm0001280

More Laws for Pauses: Replication and Generalization

Taylor M. Mezaraups and David L. Gilden
Department of Psychology, University of Texas at Austin

The basic timescales governing animal life are generally determined by body size. Pauses in naturally occurring human speech were investigated to determine if pause timescales are also sensitive to body size. Reported is an analysis of pause duration allometry in recorded interviews of 61 athletes. Pauses were divided into three classes based on whether they occurred during fluid speech or whether they preceded or followed a filled pause (i.e., "um"). Allometric laws relating body size to pause size were found for all three classes—larger people take longer pauses. The derived allometric exponents were used to evaluate a theory of how people experience the passage of time. The theory associates the experience of time passage with the distal flow of time through the mathematics of bounded exponential growth. Nonlinearities inherent in the theory are shown to predict, in detail, the way body size interacts with linguistic context in the deployment of pauses. The theory provides a meaningful framework for understanding how time is experienced as a felt quantity and how pauses are negotiated in everyday speech.

Keywords: speech production, speech perception, timing

Speech is an aspect of human behavior that provides unique opportunities to study the ethology of human timing. As an acoustic signal, it is characterized by the same dimensions that distinguish, say, musical instruments—pitch, timbre, and loudness. Speech, however, is not just an acoustic signal; it is a complex system that is simultaneously organized by grammar, prosody, and meaning. One aspect of this organization is that speech, like music, is phrased—segmented, and as in music, an important contributor to the construction of phrases is the sound of silence, the taking of pauses to create moments of separation and rest.[1] Although phrase producing pauses, *segmenting pauses*, have traditionally been investigated in terms of their linguistic functions, speech planning and speech recovery, in particular, they have a much deeper significance in providing a window into people's sense of time passage. In the same sense that a stop sign provides little direction about when to start driving again, the phrase boundary markers that halt speech do not specify how the moment to recommence speaking, having stopped, might be negotiated. The obvious, but key, insight is that segmenting pauses do in fact end by choice. What people are making when they choose to end a segmenting pause is not just the pause itself, but through phrasing, they are also creating emergent global properties, such as direction and momentum. We view the conduct of speech as a type of choreography, based to a large extent on how phrase-producing pauses are deployed. At the root of this choreography is just the speaker's sense of time passage; that is essentially all a speaker has at their disposal as they drop pauses in real-time expression.

The possibility that a person's sense of time passage, and consequently their pause behavior, might be rooted in the body was suggested by previous work (Gilden & Mezaraups, 2022a) that investigated the body's role in temporal integration. Temporal integration is simply the sense that time-distributed events are related—that they go together, even though they occur at different moments in time. The musical notes that form a melody or the beat train that creates the feeling of rhythmic pulse are two examples of time-based groups created by the sense of going-together. Our investigations began with the recognition that the sense of going-together is fragile and highly constrained by temporal proximity. Neighboring events that are separated by even just a couple of seconds will generally not be perceived as going-together; instead, they will be perceived as forming a succession of unrelated events (see Fraisse [1978] for one of the original articulations of these ideas). Gilden and Mezaraups (2022a) developed objective measures of the immanent sense of "going-together" for two forms of temporal integration, rhythmic pulse and the perception of long range apparent motion, and used these measures to examine whether limiting temporal integration spans scaled with body size.

In biology, body-size scaling is formalized in terms of allometric laws–power laws that relate some aspect of animal physiology, morphology, or behavior to animal size. What distinguishes allometry from correlational analysis, is that, in biological systems, the power law exponents, even though derived through regression analysis, are treated as meaningful discoveries that play key roles in constraining

---

David L. Gilden https://orcid.org/0000-0002-5820-3076

All relevant data is posted in Texas Data Repository: https://doi.org/10.18738/T8/RJKNG5 and https://doi.org/10.18738/T8/2V5ZSE.

Taylor M. Mezaraups developed the study concept and study design. Data collection was achieved with the help of a research assistant, Emma R. Marom. Data analysis was the primary responsibility of Taylor M. Mezaraups. David Gilden and Taylor M. Mezaraups both contributed to the drafting and editing of the manuscript. Both authors approved the final version of the manuscript.

Correspondence concerning this article should be addressed to David L. Gilden, Department of Psychology, University of Texas at Austin, 108 East Dean Keeton Stop A8000, Austin, TX 78712, United States. Email: dgilden@utexas.edu

[1] Speech is littered with brief pauses, in the duration range of 0.1 to about 0.2 s, that arise due to mechanical constraints in articulation. Examples of such pauses, termed *articulatory* pauses, include the brief pause at the double consonant that is inevitable in articulating "puppy," "happy," or "duct tape."

theory. Gilden and Mezaraups (2022a) showed that the time spans that disrupt time-based groups obeyed allometric laws, placing temporal integration limits into the same framework as physiological time scales—heartbeat period, respiration period, blood circulation time, animal lifetime, and so on. Insofar as it is the construction of time-based groups that makes the world a meaningful and coherent place to live in, this perspective places the body at the center of mental life.

The time spans between neighboring events that disrupt temporal grouping tend to be long, 1–1.5 s, and so are longer than the pauses typically encountered in ordinary speech. Still, we wondered if the mechanisms that create body-size scaling in group-disrupting pauses might not also create scaling, more generally, in the sense of time passage. To the extent that pause durations in speech provide a window into the sense of time passage, this is not a difficult issue to resolve. In Gilden and Mezaraups (2022b), we studied single speaker speech that was both read (reading poems and stories) and composed (describing life decisions, describing a cartoon, giving directions from a map). In every task, there was evidence of allometry, and collectively, the allometric relation explained more than 50% of the variance in duration. Nevertheless, these speech tasks had a staged and somewhat contrived nature. We felt that it would be productive to examine the replicability of pause allometry in a natural conversational setting. To this end, we looked for a collection of speech acts that were elicited in a common context and which included speakers with widely varying body sizes.[2]

A collection that meets these criteria are the recorded interviews with athletes that are available on YouTube. These conversations have a well-known common format: one or more interviewers ask questions, and the athlete responds in off-the-cuff commentary. In such interviews, the athletes do most of the talking, they do not read from a script, and they have wide latitude in what they might say. It is also the case that different sports attract individuals varying widely in height, including the extremes, and most of these values are published (although it is not uncommon for multiple sources of height reports to slightly disagree).

Method

Participants

The data analysis is based on the audio recordings of 61 athletes (36 female; 25 male) in 17 different sports, from interviews posted on YouTube. Out of the initial set of 99 recordings collected, 38 were excluded due to not being clear enough, containing multiple athletes, or becoming corrupted. Ages ranged from 16 to 44 years (*Mdn* = 26), and heights ranged from 56 to 83 in. (*Mdn* = 70). Athletes were included in the study on the basis of both the sound quality of the interview audio and the statistical requirement that we sample a wide range of heights. A list of included athletes, along with their heights, sports, and mean pause lengths, is displayed in Table 1.

Stimuli and Procedure

The audio recordings used to extract pause lengths were derived from YouTube videos of athlete interviews. The majority of these consisted of press conferences, wherein reporters asked various questions about the athletes' performances, opinions, lives, etc. Videos selected for the study had lengths ranging from 3.97 to 7.17 min (*M* = 5.36), to ensure that each athlete was talking for at least 1 min, and clips were truncated so that no athlete contributed more than 4 min of speech. The majority of interviews included relatively long answers, with participant means ranging from 6.73 to 41.25 s (*M* = 18.03, *SD* = 7.61).

Pause Extraction

Pause extraction was accomplished by converting the YouTube interviews to MP3s, using the onlineYouTube to MP3 Music Converter (YTMP3; https://ytmp3.cc), and then employing the Sound Finder Tool in Audacity open-source digital recording software (www.audacityteam.org), to mark pause beginnings and endings on the audio waveforms. There are two issues that require attention in using a rote algorithm to extract pauses. The first is setting a floor decibel level for what counts as silence. The second is setting a minimum duration for a period of silence to be counted as a segmenting pause. Both of these issues were confronted in Gilden and Mezaraups (2022b), and the choices made there were also employed here: a threshold of silence set at −28 dB within Audacity (4% of maximum resolvable signal) and a minimum pause length of 0.25 s. These settings made the pauses extracted by the Sound Finder Tool conform to the judgments of people experienced in waveform analysis. Nevertheless, there were three contexts in which the Sound Finder Tool would occasionally place terminal markers that disagreed with human judgment: truncating word endings when the sound level fell below the dB threshold, marking room sounds as speech bursts, and omitting entire words or vowels if below the dB threshold. The corrective actions taken in these cases were straightforward and are described in detail in Gilden and Mezaraups (2022b).

Each track was listened to individually, in order to check the placement of the terminal labels. During terminal verification we added additional labels: *a*, to denote athlete (not interviewer) speech, and *fp*, to denote speech bursts that were utterances of a filled pause.[3] Once we were satisfied that all labels were correct, the labeled track was exported as a text file, and pause durations were computed as gaps between ending and beginning terminals. Duration outliers longer than 1.75 s were removed, in view of the leverage that outliers have in highly skewed timing distributions. Regardless, the duration distributions were quite sparse beyond 1.75 s, containing fewer than 0.5% of pauses.

Filled pauses were specifically marked, because they may have important roles to play in how pauses both immediately before and after are deployed by the speaker and interpreted by the listener

---

[2] While body size could be operationalized in many ways, we typically use height. While mass is the common power law base in the field of allometry, it is problematic to use in people. Actual mass in people is highly variable, because the mass of adipose tissue is variable; it varies with any number of factors that are not relevant to allometry, including socioeconomic status, lifestyle choices, health issues, age, and zip code. A better metric, fat-free mass, requires specific equipment to measure and would not have been possible to obtain in our sample of famous athletes. Furthermore, it has been known since Quetelet (1842) that adult human weight nominally scales as height squared (see Heymsfield et al., 2007 for empirical data), making it simple to approximate a mass-specific law using easily accessible height information.

[3] Filled pauses are not actually pauses (the term is unfortunate but established), but rather meaningful utterances, like "um," "uh," or "er," that often acknowledge a long pause is in progress (or about to be) and that more speech is forthcoming (Clark & Fox Tree, 2002; Maclay & Osgood, 1959; Swerts, 1998).

Table 1
Athlete Demographic Information and Mean Pause Length

Athlete	Height (in.)	Sport	Mean pause (s)
Ailing Eileen Gu	66	Skiing	0.46
Alec Yoder	68	Gymnastics	0.53
Ali Aguilar	67	Softball	0.53
Amanda Kessel	65	Hockey	0.65
Amanda Zahui B.	77	Basketball	0.76
Antonee Robinson	72	Soccer	0.63
Ariel Torres	68	Karate	0.53
Bianca Belair	67	Wrestling	0.41
Blake Griffin	81	Basketball	0.63
Breanna Stewart	76	Basketball	0.50
Brian Irr	76	Karate	0.65
Chloe Kim	63	Snowboarding	0.70
Cody Bellinger	76	Baseball	0.78
Cody Rhodes	73	Wrestling	0.66
Dana Rettke	80	Volleyball	0.51
Dearica Hamby	75	Basketball	0.61
Delaney Spaulding	67	Softball	0.54
Deonte Harty	66	Football	0.62
D. J. Augustin	71	Basketball	0.53
Elisa Au	66	Karate	0.48
Emma Raducanu	69	Tennis	0.62
Erriyon Knighton	75	Track	0.80
Freddie Freeman	77	Baseball	0.64
Gio Reyna	73	Soccer	0.69
Hali Flickinger	65	Swimming	0.56
Hilary Knight	71	Hockey	0.50
Irad Ortiz Jr.	63	Jockey	0.54
Isaiah Thomas	69	Basketball	0.70
J. J. Taylor	66	Football	0.72
Jade Carey	61	Gymnastics	0.52
Jose Altuve	66	Baseball	0.51
Karen Chen	61	Figure skater	0.61
Karl-Anthony Towns	83	Basketball	0.61
Katie Grimes	70	Swimming	0.80
Katie Ledecky	72	Swimming	0.60
Kayla Caffe	72	Volleyball	0.56
Kelsie Whitmore	66	Baseball	0.53
Kevin Durant	82	Basketball	0.90
Kyrie Irving	74	Basketball	0.62
Lauren Stivrins	76	Volleyball	0.60
Lindsey Jacobellis	65	Snowboarding	0.56
Liz Cambage	80	Basketball	0.66
Madison Lilley	71	Volleyball	0.55
Margaret Purce	65	Soccer	0.61
Maria Sakkari	68	Tennis	0.82
Mariah Bell	64	Figure skater	0.49
Michael Kemerer	69	Wrestling	0.58
Mikaela Shiffrin	67	Skiing	0.59
Monica Abbott	75	Softball	0.64
Natasha Cloud	72	Basketball	0.46
Nathan Chen	65	Figure skater	0.50
Rachael Blackmore	66	Jockey	0.69
Raevyn Rogers	69	Track	0.49
Roger Federer	73	Tennis	0.66
Sakura Kokumai	60	Karate	0.61
Simone Biles	56	Gymnastics	0.48
Sue Bird	69	Basketball	0.56
Tom Brady	76	Football	0.64
Trae Young	73	Basketball	0.53
Vincent Zhou	69	Figure skater	0.51
Zoi Sadowski-Synnott	65	Snowboarding	0.57

(Clark & Fox Tree, 2002). A filled pause, for example, might indicate the acknowledgment that a delay is in progress and that speech is forthcoming (Maclay & Osgood, 1959). This context was deliberately created in Gilden and Mezaraups (2022b), where we asked participants difficult questions such as, "Can robots create art?." In the present study, we only marked filled pauses that occurred during the fluent speech that formed athletes' answers to reporters' questions. Pauses between reporter questions and athlete responses were not included, as experience with the video recordings indicated that this time period was highly variable and sensitive to situational factors that were unrelated to language process. In the interview context, maintaining the floor may be less of an issue, and the filled pauses taken during fluent speech may be emphasizing discourse structure (Swerts, 1998), or perhaps signaling that a delay is going to occur, both of which may help to increase comprehension (Brennan & Williams, 1995). Filled pauses within fluent speech created two additional pause classes: pauses before filled pauses and pauses after filled pauses. Segmenting pauses, to be clear, are all pauses exceeding 0.25 s in length that are not associated with a filled pause.

An example of how the audio track is partitioned by the Sounder Finder Tool is shown in Figure 1. This particular recording is from Simone Biles, the shortest person in our sample. The words contained in the audio file are written above the signal, and the pauses are marked. In this brief sample, there is an ending of a speech burst, a beginning of a speech burst, and a filled pause, ("um"), all of which lead to the extraction of a before-fill pause and an after-fill pause.

The Biles snippet reveals some of the typical pause behavior we observed throughout the interviews. In the second speech burst, for example, there is a natural place for a pause after "variables," where a comma would be placed in a written transcription. Simone Biles does not generate a pause here of length >0.25 s to be picked up as a segmenting pause. A taller person might take a longer pause in this context, exceeding the 0.25 s cut-off, and so breaking the second speech burst up into two bursts or more. It was generally the case that shorter athletes had fewer markable pauses within their responses and consequently displayed longer speech bursts, $r(59) = -.25$, $p = .03$. Both raw and aggregate data is posted in Texas Data Repository: https://doi.org/10.18738/T8/RJKNG5 and https://doi.org/10.18738/T8/2V5ZSE.

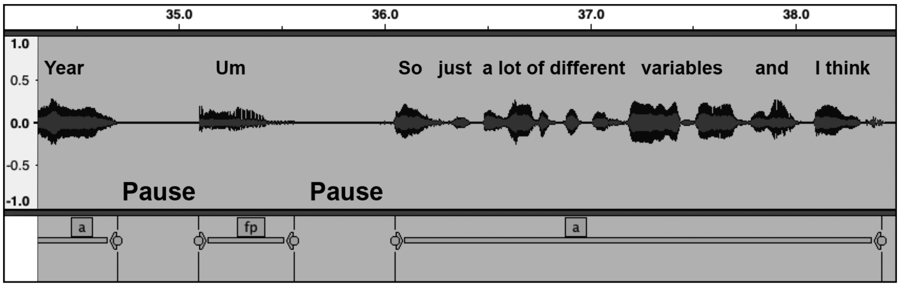
Results

Overall, the data analyzed in this study consisted of 4,110 pauses, taken over the course of 5.4 hr of recorded speech. Table 2 gives basic statistical information on the pause counts and duration distributions. The first column of Table 2 gives the total number of pauses in each pause class, Columns 2 and 3 give participant averaged statistics for pause count and time spent pausing, and the last three columns give distribution moments for each pause class, pooled over participants. The vast majority of pauses taken, by an order of magnitude, were in the segmenting class (not bounded by a filled pause)—not surprising given the fact that speech is generally fluid and not littered with "ums" and "ers." These distributions are illustrated in Figure 2.

The focus of this study was on body-size scaling, and these results are presented in Figure 2, as regressions of mean pause duration against participant height for each pause class. The regressions were computed in the log–log plane, as allometric laws are typically expressed as power laws of mass or body size; animal property $\sim (\text{body size})^b$. In the log–log plane, the exponent, b, is easily extracted in a simple regression as the slope. In computing regression

Figure 1

Screenshot of a Speech Waveform in Audacity, With Accompanying Words From the Interview of Simone Biles



Note. Also shown are labels marking the speech burst terminals, the terminals of a filled pause, and the gaps in between terminals marked as pauses.

models within the three pause classes, five outliers were flagged for removal using the criterion of four times the mean Cook's distance (two from segmenting, one from before fill, and two from after fill). Figure 2 makes the statistical case that pause durations generally satisfy allometries. Across pause classes, the power law exponents ranged between 0.81 and 1.86, all exponents were significantly different from zero ($p \leq .001$), and the proportions of variance explained by the regressions ranged between 15% and 19%.

The partitioning of pauses into three classes was not expected to have consequence beyond the verification that filled pauses are often associated with relatively longer pauses in their immediate vicinity. However, this partition led to unanticipated findings of additional structure in the way body size impacts pause duration. Several inequalities were discovered that pose critical tests for theories of human timing that lead to pause allometry. The first finding was that the ordering of distribution means across classes (positive linear trend in M; $z = 6.7, p < .001$) was reflected in the ordering of the allometry exponents across classes (positive linear trend in b; $z = 1.75$, $p = .04$):

$$\begin{aligned}&\text{both } M(\text{segmenting}) < M(\text{before fill}) < M(\text{after fill})\\ &\text{and } b(\text{segmenting}) < b(\text{before fill}) < b(\text{after fill}).\end{aligned} \tag{1}$$

Apparently, classes with longer pauses have greater body size differentiation—steeper height regressions. The increase in exponent with increase in mean class pause duration was also found in the second experiment of Gilden and Mezaraups (2022b), where we measured long pauses following difficult questions. However, with the greater resolution provided by the present study, we recognized that the exponents were also quite sensitive to class mean. A 25% increase in mean between segmenting and after-fill pauses led to a factor of 2 increase in exponent. This sensitivity suggests that a theory of human timing that can accommodate these inequalities be built around an underlying nonlinearity in the relationship between pause duration and body size.

Table 2

Descriptive Statistics for Three Pause Classes

Pause class	*n*	Pause count *M*	Pause time *M*	Pause duration distribution moments *M*	*SD*	Skew
Segmenting	3,505	57.5	32.3	0.56	0.26	1.62
Before fill	349	6.23	3.96	0.63	0.26	0.99
After fill	256	4.74	3.36	0.71	0.35	0.79

The second finding was that, while pauses both before and after filled pauses were lengthened as expected, the degree of lengthening interacted with body size. Table 3 makes it clear that pause class affected people of different heights differently; as height increased, so did the dispersion of pause duration means among the three pause classes. This was not a small effect; in the shortest group, there was no measurable distinction in mean pause duration between segmenting pauses and pauses associated with fills. This finding was most unexpected, in view of the semantic meanings that filled pauses are theorized to convey. Again, a theory of pause allometry should be able to account for this somewhat odd outcome.

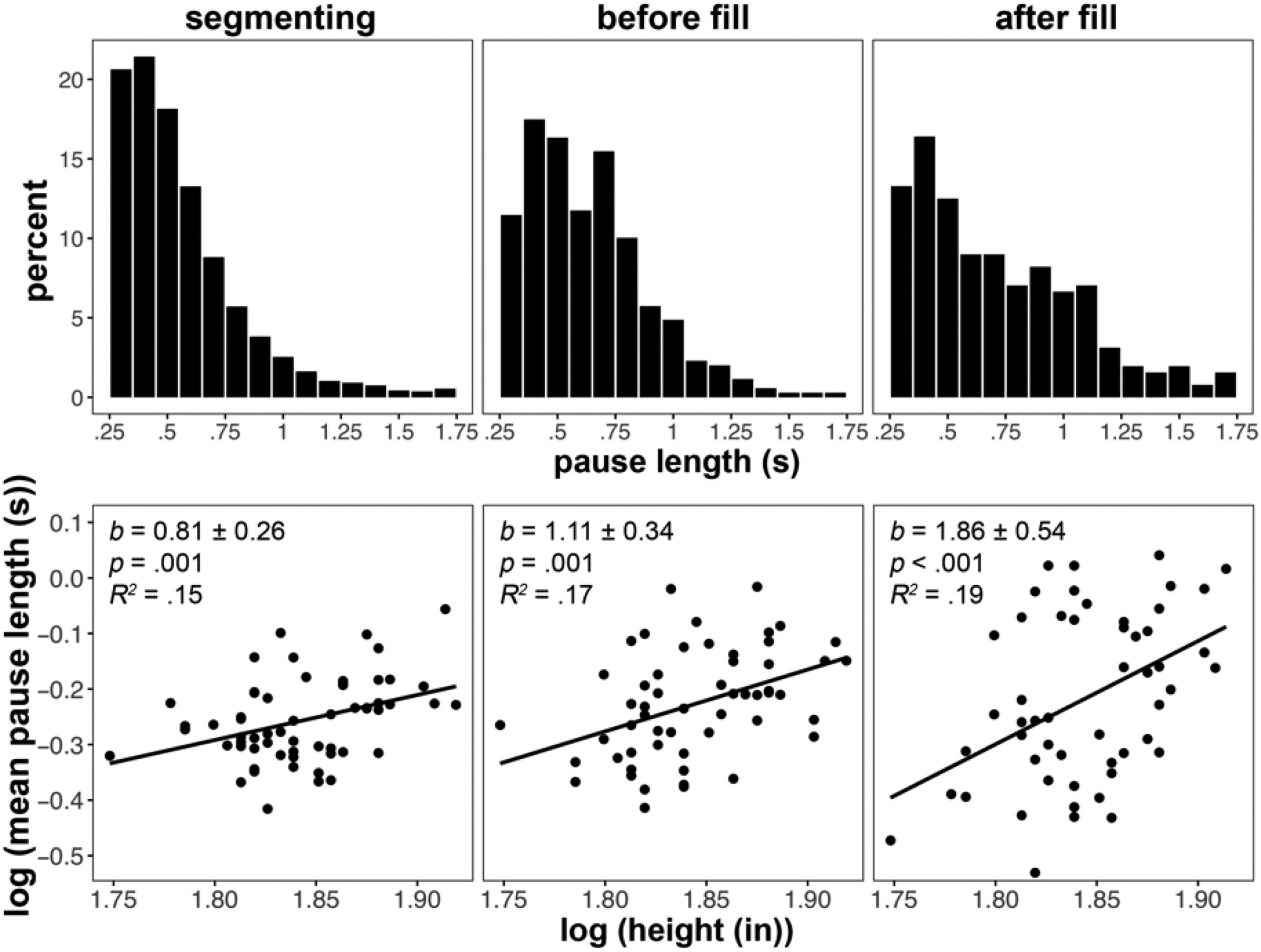
A particularly salient result from this study is that, regardless of pause class and the context in which a pause is deployed, pauses rarely exceeded 1.5 s (<2%). This finding is reiterated throughout the pause duration literature (see e.g., Campione & Véronis, 2002), evidence in itself that 1.5 s may be a landmark in the ethology of human timing. The value of 1.5 s is not to be taken as a fixed constant such as might be encountered in physics, but as a narrow region, perhaps better notated as 2 ± 1 s. The notion that 2 ± 1 s is a watershed in the experience of temporal duration derives from other sources as well. One source is psychophysical evidence from studies of duration discrimination (Getty, 1975; Grondin et al., 1999) that short durations, less than about 1.2 s, are "perceived" or "felt" (Tomassini, 2016), while longer durations require explicit estimation to be reckoned. More generally, 2 ± 1 s acts as a kind of proximity constraint for the formation of temporal groups and scenes (Gilden & Mezaraups, 2022a). Successive events that are separated by less than 2 ± 1 s tend be perceived in relation to another and so as belonging to a common group or scene. Successive events separated by more than 2 ± 1 s tend to be perceived simply as a succession of unrelated events. In this way, a pause of 2 ± 1 s effectively acts as scene break within a speech event.

Discussion

The central result is also the simplest to state: that we have replicated the finding of allometry in three separate classes of pause production—in segmenting pauses in fluid speech, and in two classes of pauses associated with filled pauses. These replications in athlete

Figure 2

Pause Distributions and Mean Pause Length Regressions on Height for the Three Pause Classes, in the Log(Inches) − Log(Seconds) Plane



speech are ethologically distinct from the in-laboratory speech acts recorded for Gilden and Mezaraups (2022b), in that interview speech is conversational and unconstrained. The replication represents independent confirmation that allometry in speech pauses is a real and robust phenomenon. A second notable result was that, by partitioning pauses into three classes, we discovered coordinate inequalities in the allometric exponents and class means that challenges a theory of pause allometry we proposed in Gilden and Mezaraups (2022b).

In this theory, the core construct is the notion of felt time, and the core idea is that, when a speaker takes a pause, there is a feeling of pause fullness that grows from the moment of pause initiation to a point where the pause duration feels, literally, sufficiently full to recommence speech. What counts as an appropriate level of pause fullness will depend upon a myriad of factors, including the type of linguistic boundary that initiated the pause, the demands of speech planning and speech recovery, the cadence of speech surrounding the pause, and finally, whatever semantic meaning is conveyed by the pause. The general shape of a pause fullness function is constrained by the requirement that pause fullness saturate at around 1.5 s of growth, where further temporal discrimination on the basis of felt time should not be possible. Consequently, a pause fullness growth function must asymptote at about 1.5 s, as illustrated in Figure 3 (reprinted from Gilden & Mezaraups, 2022b).

Table 3

Mean Lengths of Three Pause Classes for Three Height Groups

Pause class	Short (<67 in.)	Medium (67–74 in.)	Tall (>74 in.)
Segmenting	0.54 (0.02)	0.54 (0.02)	0.62 (0.02)
Before fill	0.54 (0.03)	0.61 (0.03)	0.69 (0.03)
After fill	0.54 (0.05)	0.64 (0.05)	0.77 (0.05)

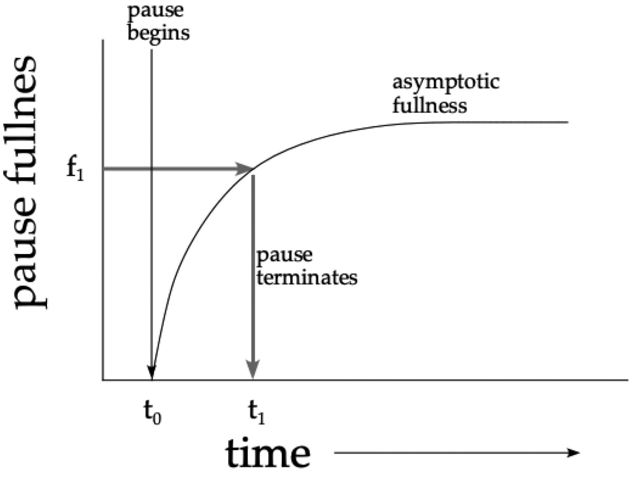
Note. Standard error in parentheses.

A formal expression for bounded pause fullness growth may be constructed through generic models of sensory integration. A model that leads to the bounded growth illustrated in Figure 3 is the leaky integrator (see Gilden & Mezaraups, 2022b; Toso et al., 2021), which has the solution under the assumption of a constant supply rate:

$$f(t) = f_a(1 - e^{-t/\tau}), \tag{2}$$

Figure 3

Anatomy of a Pause



Note. Pause onset at t_0 initializes an epoch of sensing the fullness of elapsed time. The pause ends at t_1, when the process arrives at the state of fullness f_1. Reprinted from “Laws for Pauses,” by D. L. Gilden and T. M. Mezaraups, 2022b, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48(1)*, pp. 142–158 (https://doi.org/10.1037/xlm0001103). Copyright 2022 by the American Psychological Association.

where f_a is the asymptotic level of fullness, and τ is a growth time scale. The benefit of dressing the informal idea of growing pause fullness in a mathematical expression is that it identifies a time scale, τ, as a portal for allometry to enter language behavior.

τ is an interesting object from the point of view of psychological theory. It is a definite time, but it is also a time scale that is a function of system parameters (see Gilden & Mezaraups, 2022a for a more detailed discussion of time scales and their parametric dependencies). The system that τ is embedded in—a speaker—is a biological system, but it is performing within a psychological-social context–speech. τ is a clearly a construct of potentially formidable complexity, yet it is the case that time scales in biological systems generally do satisfy allometries (a famous example is the Kleiber Law). It is also the case that if τ satisfies an allometry, then granted two structural assumptions, so too will pause duration. First, if τ allometry is typical of living systems, then it will be an increasing function of body size. The mathematics of the exponential then requires that shorter people with shorter fullness growth times reach any given level of pause fullness earlier than taller people with longer fullness growth times. Secondly, if in any given linguistic situation, people terminate their pauses at common levels of fullness, shorter people will terminate their pauses earlier than taller people. In this way, τ allometry ends up being realized as pause duration allometry.

The logic of pause allometry is depicted in Figure 4. For the purposes of illustration, three growth functions are plotted, corresponding to three values of the growth timescale, τ. Also shown are three generic levels of pause fullness that are suggested by the mean durations found for the different classes. As $M(\text{segmenting}) < M(\text{before fill}) < M(\text{after fill})$, and because the pause fullness growth functions monotonically increase with elapsed time, the pause fullness levels typical of the three classes will satisfy $f_s < f_{bf} < f_{af}$. Allometry is created in each pause class as a simple consequence of the circumstance that steeper growth functions (smaller body size, smaller τ) arrive at a given level of pause fullness earlier than the shallower growth functions (larger body size, larger τ). Here we have attempted to depict the effect of body size scaling through the visual heuristic of segment length between the filled dots in each pause class. Segment length provides a visual measure of the degree to which body size affects pause duration in a given pause class, and so is closely related to the class allometric exponent.

Figure 4

Pause Allometry in Three Classes of Pauses That Are Initiated in Three Distinct Regions of Pause Fullness: Segmenting Pauses at f_s, Pauses Before Fills at f_{bf}, and Pauses After Fills at f_{af}

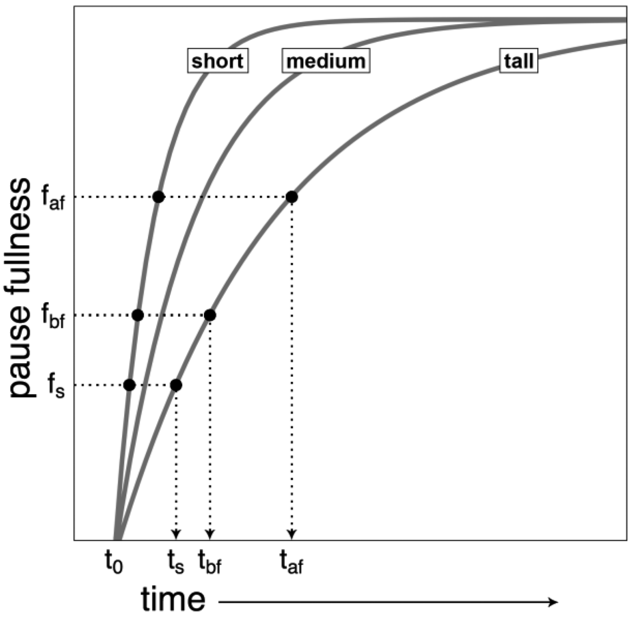


Figure 4 also illustrates how a bounded exponential pause fullness growth function generates exactly the nonlinearities required to understand both (a) the sensitivity of exponent to mean class duration and (b) the interaction between body size and the degree of pause lengthening induced by filled pauses. Nonlinearity in exponent growth is illustrated by the rapid rate at which segment lengths between the filled dots grow as the triggering fullness level moves upward on the *y*-axis. In essence, nonlinear exponent growth with pause fullness is a consequence of the empirical fact that felt time has a boundary, and so pause fullness functions flatten. The shape of the pause fullness functions also explains why average pause times in the three classes displayed in Table 3 are more disparate in taller people. To illustrate this, we have explicitly drawn the typical pause times that the tallest people would realize in the three pause classes. The pause times explicitly marked on the *x*-axis (t_s, t_{bf}, t_{af}) are intended to illustrate the entries in the third column of Table 3 (class means for tall people). Focusing now on the medium size people, not explicitly illustrated, at equivalent levels of pause fullness, their pause times are more similar across pause classes, because their growth function is relatively steeper. And in the shortest group, also not illustrated, the pause times are even more compressed across classes, as their function is even steeper. These trends produce the data structure of Table 3 in some detail. That this simple theory is able to explain both of these rather subtle aspects of the athlete pause data is not trivial, especially in view of the circumstance that the theory was minimally constructed in Gilden and Mezaraups (2022b), only to explain how variation in body size might produce variation in pause duration. Recognizing that the notion of felt time is fairly fuzzy, the theory is not fuzzy and appears to capture in meaningful ways how people experience the passage of time.

Any formal psychological theory that is not meaningfully based in biology or neuroscience must be taken as it is offered—as a mathematical construction that hopefully makes meaningful contact with observed data. Confidence in the model is inextricably linked to the extent to which the model has been exposed to observed data and to the possibility that it is not in agreement with observed data. In this spirit, we are interested how this model might be further tested. Insofar as the model generally predicts allometry, any circumstance where pause durations do not scale with body size would be of great interest. More specifically, because the model is constructed around a decaying exponential, it generates pause fullness growth functions that have a particular shape, and that shape itself may be exposed to further observational tests.

The particular way in which the pause fullness functions fan out in Figure 4 is where the model may be most vulnerable. The model is constructed to predict that allometric exponents will increase with mean pause length, and while this is observed in the athletes' data, it is worth exploring further. In particular, it would be interesting to see if allometry might be removed entirely in calibrated classes of relatively short pauses. Calibrated classes of short pauses might be generated by exposing people to language snippets such as "blah, blah, blah" or "yeah, yeah, yeah." Pauses with mean duration <0.5 s would be expected to populate the distributions produced by these language opportunities and to have little height sensitivity. A weakness that is inherent in testing the short pause range of speech is that there are statistical constraints on what allometry might be observed. Short pauses, being short, will not show a great deal of variability. It is a general property of reaction time data, and timing data in general, that standard deviations are roughly proportional to means, even if they do not satisfy a strict scaling law. Where there is low variability, there is little opportunity for any factor, including height, to capture substantial amounts of variance.

Exposing the model to data with long pause durations might be more productive. The model predicts that pause fullness saturates for all people and that the approach to saturation is very sensitive to body size. The asymptotic region is a large target for testing the model, but it also comes with methodological issues. The first issue is that, in natural speech, it is rare for people to take long pauses. The distributions shown in Figure 2 make this clear. The long duration portion of the pause distribution is also not pure, in the sense that it is not populated by a single process. There are certainly long pauses that are ended by the sense that now is a good time to recommence speech—that is, by a decisional process, but there are also long pauses that reflect dysfluencies and distraction. There may be better ways to probe the asymptotic region of the fullness function than through speech analysis. The model, to be clear, in no way is limited to speech pauses, and is intended to be a general model of the time sense.

A potentially fertile way of probing the rate at which fullness functions asymptote with height is suggested by Grondin et al. (1999). They measured the point at which counting helps in discriminating the lengths of time intervals. Using a clever methodology, they were able to show that there is a specific point, around 1.2 s, where counting helps. The interpretation is that, for shorter times, people have a sense of pause (time interval) fullness, and they can effectively use fullness in a discrimination task. For longer times, the fullness function has saturated, and then counting helps in discriminating time intervals. This methodology might be useful in determining if there is allometry in the point at which counting helps. In this way, a direct analysis of allometry in the approach to pause fullness saturation might be possible.

The Grondin et al. methodology is also relevant, because it raises the question of what is counting. Counting is a way of bridging an interval beyond the reach of pause fullness, by chopping the interval up into bits that are on the pause fullness function. An interesting study might be an investigation into whether there is allometry simply in the rate at which people count. The model does predict that, if count intervals feel the same to all people, then there should be allometry in the distal count rate.

References

Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, *34*(3), 383–398. https://doi.org/10.1006/jmla.1995.1017

Campione, E., & Véronis, J. (2002). A large-scale multilingual study of silent pause duration. In B. Bel & I. Marlien (Eds.), *Proceedings of speech prosody 2002* (pp. 199–202). Laboratoire Parole et Langage.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*(1), 73–111. https://doi.org/10.1016/S0010-0277(02)00017-3

Fraisse, P. (1978). Time and rhythm perception. In E. Carterette & M. Friedman (Eds.), *Handbook of perception* (Vol. VIII, pp. 203–254). Academic Press.

Getty, D. (1975). Discrimination of short temporal intervals: A comparison of two models. *Perception & Psychophysics*, *18*(1), 1–8. https://doi.org/10.3758/BF03199358

Gilden, D. L., & Mezaraups, T. M. (2022a). Allometric scaling laws for temporal proximity in perceptual organization. *Psychological Review*, *129*(3), 457–483. https://doi.org/10.1037/rev0000307

Gilden, D. L., & Mezaraups, T. M. (2022b). Laws for pauses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(1), 142–158. https://doi.org/10.1037/xlm0001103

Grondin, S., Meilleur-Wells, G., & Lachance, R. (1999). When to start explicit counting in a time-intervals discrimination task: A critical point in the timing process of humans. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 993–1004. https://doi.org/10.1037/0096-1523.25.4.993

Heymsfield, S. B., Gallagher, D., Mayer, L., Beetsch, J., & Pietrobelli, A. (2007). Scaling of human body composition to stature: New insights into body mass index. *The American Journal of Clinical Nutrition*, *86*(1), 82–91. https://doi.org/10.1093/ajcn/86.1.82

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, *15*(1), 19–44. https://doi.org/10.1080/00437956.1959.11659682

Quetelet, A. (1842). *A treatise on man and the development of his faculties: Now first translated into English*. William and Robert Chambers.

Swerts, M. (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics*, *30*(4), 485–496. https://doi.org/10.1016/S0378-2166(98)00014-9

Tomassini, A. (2016). Wittmann, M. Felt time—The science of how we experience time, MIT Press. *Perception*, *46*(11), 1333–1334. https://doi.org/10.1177/0301006617718993

Toso, A., Fassihi, A., Paz, L., Pulecchi, F., & Diamond, M. E. (2021). A sensory integration account for time perception. *PLoS Computational Biology*, *17*(1), Article e1008668. https://doi.org/10.1371/journal.pcbi.1008668

Received February 22, 2023
Revision received June 11, 2023
Accepted June 22, 2023 ■