Cerebellar activity and stuttering: Comments on Max and Yudman (2003)

Peter Howell
University College London

Wing and Kristofferson (1973) introduced a simple motor control task to study timing of actions. In the original task, participants tap in time to an isochronous metronome. After the experimenter judges the participant has settled into the responses, the metronome ceases, and the participant attempts to continue unaided at the metronome rate. After cessation of the metronome, the overall variability of the responses is obtained. The overall variability is then decomposed into two variance components that reflect, according to Wing and Kristofferson, that due to an external timekeeper and that due to timing variability associated with implementation of the motor responses. The modeling assumption made to obtain implementation variance is that motor processes that lead to one response occurring earlier than intended will be compensated for in the next interval. Consequently, the lag one autocorrelation can be used to estimate how implementation variance affects these two intervals. When the contribution of this variance component is subtracted from the total variance, what is left is the variance that reflects the operation of the external timekeeper. Ivry (1997) has shown that the two variance components are affected by lesions to different parts of the cerebellum. Medial damage affects implementation variance and lateral damage affects external timekeeper variance. Though there is little disagreement about the role of the cerebellum in motor timing (Max & Yudman, 2003), this does not mean that these are the only roles performed by the cerebellum (see, for example, Ohyama, Nores, Murphy, & Mauk, 2003 for another proposed function of the lateral cerebellum). It should also be cautioned that establishing neural region function from lesion studies is not straightforward and it is essential to conduct non-lesion studies as well (Ohyama et al., 2003).

Stuttering is a disorder that affects speech output timing (Perkins, Kent & Curlee, 1991; Starkweather, 1985; Wingate, 1976) and can be ameliorated if speech timing is voluntarily (Barber, 1940) or involuntarily changed by, for example, delayed auditory feedback (DAF) (Goldiamond, 1965). This work suggests that timing control may be associated with the disorder that could involve cerebellar systems. To test for a cerebellar role, Howell, Au-Yeung and Rustin (1997) obtained data from children who stutter and fluent controls in a version of the Wing and Kristofferson task. The task they employed involved tracking the movement of a sinusoidally-moving visual target by moving the lower lip. They reported that implementation variance was higher in speakers who stutter than in controls. A related study in adult speakers who stutter, using a different task and analysis procedure, also found higher variability than in controls (Boutsen, Brutten & Watts, 2000).

Max and Yudman (2003), using different variants on the Wing and Kristofferson task (bilabial contact in speech and non-speech tasks, and thumb-index finger contact in a finger movement task), failed to find timing differences between adults who stutter and fluent controls in either variance component. From this, they argued that stuttering is not

Address Correspondence to: Peter Howell, Department of Psychology, University College, London, Gower St., London WC1E 6BT, England. Telephone Number: +44 207 679 7566, Fax number +44 207 436 4276, Email address: p.howell@ucl.ac.uk.

associated with the specific cerebellar mechanisms subserving isochronous rhythmic timing. The dismissal of cerebellar timing processes seems premature given the consistent evidence from Boutsen et al. (2000) and Howell et al. (1997) and the fact that imaging studies have shown cerebellar activity is associated with the disorder in adults (de Nil, Kroll & Houle, 2001; Ingham, Fox & Ingham, 1997). There are also a number of theoretical models of stuttering (Howell, 2002; Neilson & Neilson, 1991; Nudelman, Herbrich, Hoyt & Rosenfield, 1987) and empirical findings on tracking behavior (Zebrowski, Moon & Robin, 1997) that are also consistent with cerebellar involvement in the disorder. Moreover, there are some reasons why the Max and Yudman study could have failed to find differences between their fluency groups.

The first reason concerns the procedures employed. Their participants were given a fixed number of entrainment trials where they heard the sound, after which the training sequence was switched off. The usual procedure in the Wing-Kristofferson task, as indicated earlier, is for the experimenter to judge when the participant is entrained in his or her productions to the sequence played and then switch the sequence off. No such check appears to have been made in Max and Yudman's (2003) study as to whether the participants were entrained or not before data acquisition in the continuation phase commenced. The question this raises is whether the Max and Yudman procedure provides data that are appropriate to estimate implementation and timekeeper variance. According to the Wing-Kristofferson model, until participants are entrained, their clock is not set to the required rate, and until they are settled in to this rate, the compensation process that leads to implementation variance cannot be estimated. The size of the variance estimates during entrainment is not a reflection of how close performance is to that the participant achieves after entrainment is complete. Rather, the variance estimates obtained are not meaningful until the participant has been entrained into the task. To illustrate, assume a participant starts at a rate that does not correspond with the metronome (to describe this in Wing and Kristofferson's terms, having set the clock going at the wrong rate). In this situation, he or she would adjust the timekeeper to the required rate (slowing if the rate is too fast, speeding up if the rate is too slow), and any over-adjustments would require a compensatory change in the clock setting. These timekeeper adjustments can mimic the effects associated with implementation variance, which would inflate this variance estimate and reduce the timekeeper estimate. Thus, it is not appropriate to use the Wing-Kristofferson procedure to obtain the variance estimates during entrainment while these adjustments are happening. A partial solution would be to drop the entrainment data from the statistical analyses (as is customary in tasks based on Wing and Kristofferson, 1973) as the variance measures would then provide a rough estimate of the processes in the Wing and Kristofferson model that they are supposed to reflect.

The second point (that also indicates why the solution just proposed is only partial) concerns the effect that including the entrainment data could have on statistical analysis. According to the above argument, it is misleading to include variance estimates obtained in the entrainment phase (that do not reflect the processes Wing and Kristofferson describe) along with the estimates obtained during the continuation phase (that do reflect the processes Wing and Kristofferson describe) when seeking to establish whether, for instance, these estimates differ across fluency groups. As argued earlier, indexes of underlying processes that can be estimated after entrainment, are not obtained by applying Wing and Kristofferson's analysis procedure to data from the entrainment phase. Including the estimates from the entrainment phase (before the participant has settled) in the statistical analysis would add noise and could lead to differences between speaker groups not being detected. Thus, the Max and Yudman analysis procedure is unusual and, arguably, misleading in terms of the impact it would have on statistical analyses.

The third (relatively minor) point follows from the previous issues discussed. The view can be raised that people who stutter have higher variance estimates in the continuation phase because they are less entrained than fluent speakers. If so, Max and Yudman's conditions would be set to increase the likelihood of detecting a difference between fluency groups and, as none was found, this could be considered to bolster their conclusion. However, if speakers who stutter are not entrained by the time they commence the continuation phase, the variance estimates are not meaningful (rather than indications of poor timing performance) for the same reasons as given when considering the entrainment phase. Thus, if the argument that speakers who stutter are more likely to start the continuation phase before entrainment is complete is correct, using a fixed length sequence invalidates group comparisons.

The fourth comment applies to Max and Yudman's speech and non-speech tasks alone. In these conditions in Max and Yudman's study, the participants were instructed to make their responses in synchrony with acoustic events. As the authors state: "In the speech task, the response consisted of the syllable /pa/. In the orofacial nonspeech task, the response was a 'popping' sound produced by bilabial closing and opening movements while slightly reducing, rather than increasing, oral air pressure." If participants are asked to make an acoustic response, then these responses should be used in analysis. Similarly, if participants are required to produce a tracking response with the lower lip, then movement of that articulator should be used as a measure (as in Howell et al., 1997). Despite the fact that acoustic responses were required, acoustic measures were not used by Max and Yudman (2003) for their analysis; instead they used an articulatory measure (which has the advantage that it may relate to the measure they obtained in their finger movement task). Analysis using acoustic onset would not be expected to give the same results as when using articulatory onsets. The basis of this is the known fact that when sequences of sounds including different syllables are spoken isochronously, the acoustic onsets are not evenly spaced in time (the p-center effect) (Fowler, 1979; Morton, Marcus & Frankish, 1976). One explanation that has been proposed is that all speakers time articulatory gestures to be isochronous that leads to acoustic onsets being anisochronous (Fowler, 1979). Max and Yudman's speech and non-speech tasks required isochronous production of the <u>same</u> syllable. The Wing-Kristofferson analyses on homogenous sequences would be expected to give the same results whether acoustic or articulatory markers of isochrony are used if the discrepancy between the two is constant across speakers for that syllable (the markers would than just have a constant offset that would not affect variances). However, individual differences across speakers have been reported when acoustic onsets were measured (Scott & Howell, 1992). If the articulatory markers that are supposed to be fixed for different utterances (Fowler, 1979) also remain constant across speakers, the variation between speakers (Scott & Howell, 1992) would be specifically associated with the acoustic markers. Whether this variation depends on fluency group and when variance is divided into implementation and timekeeper components are an empirical questions. The two implications are: 1) An analysis of the Max and Yudman data using acoustic onsets, would be more appropriate than their own analysis which uses articulatory measures. 2) An analysis using acoustic onsets might well produce different results from those reported by Max and Yudman (2003).

Fifth, the markers for each event in a sequence were computed automatically by Max and Yudman (2003). They combined measures over lower lip, upper lip and jaw to obtain a lip aperture (LA) measure in the speech and non-speech tasks. In our experience (Howell & Sackin, 2002), small fluctuations in estimates (that arise when parameters are extracted automatically) can have marked effects on variability estimates in variants of the Wing and Kristofferson task. This is likely to be a particular problem with the LA parameter as it is made up from different contributions from movement of the upper lip, lower lip and jaw

that, in the case of jaw and lower lip movement, correlate. While a case can be made for using a combined measure across the articulators, more indication about the relation of this measure to movement of the individual articulators and acoustic onsets (the response they required in their procedure) is necessary. For instance, how does the filtering and subsequent combination of the individual articulatory signals affect the timing response of the resulting signal, and does the combined response lead to greater discrepancies relative to acoustic onset to those noted for individual articulators (p-center effect)? Without this information, it is difficult to compare studies that use different articulatory responses and to see whether the data preparation has led to Max and Yudman's (2003) null results.

Though Wing and Kristofferson represent their model as a central timekeeper on which implementation variance is overlaid, the two components could represent independent cerebellar processes that are activated in different ways. Nothing in the estimation procedure precludes this possibility (Wing, personal communication). Implementation variability would occur all the time. Elsewhere we have argued that the external timekeeper is activated in prescribed circumstances, for example when precise timing is called for, as in the Wing-Kristofferson task or when DAF is switched on (Howell & Sackin, 2002). In their closing remarks, Max and Yudman seem to want to dismiss external timekeeping as problematic for people who stutter, and to favor an intra- or inter-gestural timing perspective. However, even if this is accepted, it is possible that implementation variability might represent an intra-gestural timing component, located in the cerebellum and involved in organizing motor representations for output (Howell, 2002; Howell & Au-Yeung, 2002).

Overall, there are many unanswered questions about the Max and Yudman procedure: Were the participants properly entrained? Does the inclusion of data from the entrainment phase add noise to their analysis? Should an articulatory response have been used, given the responses they asked participants to make, and would different results have been obtained if they had used acoustic measures? What does their data preparation do to LA signal timing? These questions raise doubts about Max and Yuderman's (2003) rejection of the view that cerebellar timing mechanisms are implicated in stuttering. They seem to regard dismissal of cerebellar mechanisms in the disorder as essential to their intra- and inter-gestural mechanisms interpretation. While external timekeeping operations may not be consistent with intra- or inter-gestural timing models, it is not clear whether this also applies to implementation timing variability associated with cerebellar structures. Implementation variance could be the basis of intra-gestural timing in the test utterances used by Howell et al. (1997), Howell & Sackin (2002) and Max and Yuderman (2003). It appears that ruling out cerebellar involvement in stuttering is not warranted.

Acknowledgments

References

Barber V. Studies in the psychology of stuttering, XVI. Rhythm as a distraction in stuttering. Journal of Speech Disorders. 1940; 5:29–42.

Boutsen FR, Brutten GJ, Watts CR. Timing and intensity variability in the metronomic speech of stuttering and nonstuttering speakers. Journal of Speech, Language and Hearing Research. 2000; 43:513–520.

De Nil L, Kroll RM, Houle S. Functional neuroimaging of cerebellar activation during single word reading and verb generation in stuttering and non-stuttering adults. Neuroscience Letters. 2001; 302:77–80. [PubMed: 11290391]

Fowler CA. "Perceptual centers" in speech production and perception. Perception & Psychophysics. 1979; 25:375–388. [PubMed: 461098]

Goldiamond, I. Stuttering and fluency as manipulatable operant response classes. In: Krasner, L.; Ullmann, LP., editors. Research in Behavior Modification. New York: Holt, Rinehart & Winston; 1965. p. 106-156.

Howell, P. The EXPLAN theory of fluency control applied to the treatment of stuttering by altered feedback and operant procedures. In: Fava, E., editor. Current Issues in Linguistic Theory series: Pathology and therapy of speech disorders. Amsterdam: John Benjamins; 2002. p. 95-118.

Howell, P.; Au-Yeung, J. The EXPLAN theory of fluency control and the diagnosis of stuttering. In: Fava, E., editor. Current Issues in Linguistic Theory series: Pathology and therapy of speech disorders. Amsterdam: John Benjamins; 2002. p. 75-94.

Howell, P.; Au-Yeung, J.; Rustin, L. Clock and motor variance in lip tracking: A comparison between children who stutter and those who do not. In: Hulstijn, W.; Peters, HFM.; van Lieshout, PHHM., editors. Speech Production: Motor Control, Brain Research and Fluency Disorders. Amsterdam: Elsevier; 1997. p. 573-578.

Howell P, Sackin S. Timing interference to speech in altered listening conditions. Journal of the Acoustical Society of America. 2002; 111:2842–2852. [PubMed: 12083218]

Ingham, RJ.; Fox, PT.; Ingham, JC. An H_2O^{15} positron emission tomography (PET) study on adults who stutter: findings and implications. In: Hulstijn, W.; Peters, HFM.; van Lieshout, PHHM., editors. Speech Production: Motor Control, Brain Research and Fluency Disorders. Amsterdam: Elsevier; 1997. p. 293-306.

Ivry, R. Cerebellar timing systems. In: Schmahmann, J., editor. The Cerebellum and Cognition. San Diego: Academic Press; 1997.

Max L, Yudman EM. Accuracy and variability of isochronous rhythmic timing access motor systems in stuttering versus nonstuttering individuals. Journal of Speech, Language and Hearing Research. 2003; 46:146–163.

Morton J, Marcus SM, Frankish CR. Perceptual centers (p-centers). Psychological Review. 1976; 83:405–408.

Neilson, MD.; Neilson, PD. Adaptive model theory of speech motor control and stuttering. In: Peters, HF.; Hulstijn, W.; Starweather, CW., editors. Speech motor control and stuttering. Amsterdam: Excerpta Medica; 1991. p. 149-156.

Nudelman, HB.; Herbrich, KE.; Hoyt, BD.; Rosenfield, DB. Dynamic characteristics of vocal frequency tracking in stutterers and nonstutterers. In: Peters, HFM.; Hulstijn, W., editors. Speech Motor Dynamics in Stuttering. New York: Springer-Verlag; 1987. p. 161-169.

Ohyama T, Nores WL, Murphy M, Mauk MD. What the cerebellum computes. Trends in Neuroscience. 2003; 26:222–227.

Perkins WH, Kent R, Curlee R. A theory of neuropsycholinguistic functions in stuttering. Journal of Speech and Hearing Research. 1991; 34:734–752. [PubMed: 1956181]

Scott, S.; Howell, P. Infinitely peak clipping speech alters its P-center. In: Auxiette, C.; Drake, C.; Gerard, C., editors. Proceedings of the Fourth International Congress of Rhythm Reception; Paris: CNRS; 1992. p. 151-156.

Starkweather, C. The development of fluency in normal children. In: Gregory, H., editor. Stuttering therapy: Prevention and intervention with children. Memphis TN: Speech Foundation of America; 1985.

Wingate, ME. Stuttering: Theory and Treatment. New York: Irvington-Wiley; 1976.

Wing AM, Kristofferson AB. Response delays and the timing of discrete motor responses. Perception & Psychophysics. 1973; 14:5–12.

Zebrowski, PM.; Moon, JB.; Robin, DA. Visuomotor tracking in children who stutter: A preliminary view. In: Hulstijn, W.; Peters, HFM.; van Lieshout, PHHM., editors. Speech Production: Motor Control, Brain Research and Fluency Disorders. Amsterdam: Elsevier; 1997. p. 579-584.