Above the Mean: Examining Variability in Behavioral and Neural Responses to Multisensory Stimuli

Sarah H. Baum [1,*], **Hans Colonius** [2], **Antonia Thelen** [1],
Cristiano Micheli [3] **and Mark T. Wallace** [1,4]

[1] Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN, USA

[2] Department of Cognitive Psychology, University of Oldenburg, Oldenburg, Germany

[3] Department of Psychology, Carl-von-Ossietzky University, Oldenburg, Germany

[4] Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

Abstract

Even when experimental conditions are kept constant, a robust and consistent finding in both behavioral and neural experiments designed to examine multisensory processing is striking variability. Although this variability has often been considered uninteresting noise (a term that is laden with strong connotations), emerging work suggests that differences in variability may be an important aspect in describing differences in performance between individuals and groups. In the current review, derived from a symposium at the 2015 International Multisensory Research Forum in Pisa, Italy, we focus on several aspects of variability as it relates to multisensory function. This effort seeks to expand our understanding of variability at levels of coding and analysis ranging from the single neuron through large networks and on to behavioral processes, and encompasses a number of the multimodal approaches that are used to evaluate and characterize multisensory processing including single-unit neurophysiology, electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and electrocorticography (ECoG).

1. Introduction

Traditional lines of neuroscientific inquiry have generally approached questions of interest by conducting group based experiments and analyses. Re-

---

* To whom correspondence should be addressed. E-mail: sarah.h.baum@vanderbilt.edu

search into multisensory function has been no exception. For example, a study designed to look at the neural correlates of a specific task in which performance on multisensory (e.g., combined auditory and visual) conditions is compared with performance on the respective unisensory conditions (e.g., auditory alone, visual alone) might look for mean differences in behavioral performance and/or neural activation patterns on multisensory *vs.* unisensory trials. Indeed, using such experimental designs, a large body of work has detailed the behavioral and perceptual benefits of having information available from multiple senses, as well as about the underlying neural networks (Alais *et al.*, 2010; Baum *et al.*, 2015; Calvert *et al.*, 2004; Murray and Wallace, 2012; Sarko *et al.*, 2012; Stein and Meredith, 1993; Stevenson *et al.*, 2014; Wallace *et al.*, 2004). More recently, there has also been an effort to understand how multisensory processes are affected by stimulus features and task instructions (Stevenson and Wallace, 2013), how they vary between subject cohorts (Baum and Beauchamp, 2014), or even within subject cohorts (Murray *et al.*, 2005; Romei *et al.*, 2013; Thelen *et al.*, 2014).

However, hidden within such group-based analyses are interesting differences across participants, as well as interesting differences within individual participants on a trial-by-trial basis. Although some of these differences are captured in classical measures of variability (such as Fano factor for neural responses), the general approach frequently neglects to focus upon inter- and intraindividual differences in variability — differences that are likely to be exceptionally meaningful in understanding the individual brain bases for performance and perception. Variability has long been studied in the fields of mathematical and physical science, and can be (and has been) readily extended to measures of cognitive and brain function. In such a framework, variability can contribute information to the system and provides a view into its inherent organizational properties. Take as a simple example the response profile of an individual neuron upon the presentation of a given sensory stimulus. Although the neuron may respond with the same mean number of spikes on successive presentations, the pattern of these spikes will likely differ on a trial-by-trial basis (Christen *et al.*, 2006). Given that information is encoded in both number and temporal patterning of activity, these differences are meaningful in regards to how that brain builds an internal representation of that stimulus. Thus, output patterns, whether they are from single neurons or large ensembles, tend to possess varying degrees of formal structure or pattern that deviates significantly from random fluctuations ('noise') and that reflect the inherent organization of the system. Critically, such patterns of self-organized variability are ubiquitous in nature and have been demonstrated in domains as diverse as condensed matter (Dutta and Horn, 1981), DNA base sequence structure (Voss, 1992), cellular automata (Ohtsuki *et al.*, 2006), and social networks (Christensen *et al.*, 1992), and have been shown to play an instrumental

role in facilitating organization and signal processing. These parallels offer an approach to the study of variability in neural and cognitive performance that take advantage of the measurement, analytic, and modeling approaches used in the physical and mathematical sciences.

As opposed to focusing only on error variance and reducing variability through statistical means (e.g., averaging, etc.), we suggest an approach that embraces these differences, and that is structured to discern how variability is an important contributor to multisensory neural encoding processes. Indeed, the role of variability in both cognitive flexibility, on the one hand, and in pathophysiological processes, on the other hand, is becoming increasingly recognized within the literature, including in the realm of healthy aging (Adamo *et al.*, 2014; Baum and Beauchamp, 2014; Beck *et al.*, 2012; Bielak *et al.*, 2010a; Di Martino *et al.*, 2008; Garrett *et al.*, 2011, 2012). In addition to decreased performance of aging subjects on a number of tasks spanning multiple domains, including memory, visuospatial abilities, and speed of information processing (Cerella and Hale, 1994; Hedden and Gabrieli, 2004; Jenkins *et al.*, 2000; Peich *et al.*, 2013), more recent work has also highlighted age-related increases in intraindividual variability on a number of these tasks (Hultsch *et al.*, 2002; Lovden *et al.*, 2013; MacDonald *et al.*, 2012; Murphy *et al.*, 2007; West *et al.*, 2002). For example, when tested on consecutive days, older adults have been shown to have more inconsistent performance on a word list learning task (Murphy *et al.*, 2007). Furthermore, changes in intraindividual variability may both precede and predict future cognitive declines (Bielak *et al.*, 2010b; Lovden *et al.*, 2007), indicating that variability may provide insight into neuronal processing beyond the average level of performance. Here we seek to bridge these questions, and to provide several perspectives on variability as seen from the multisensory viewpoint.

2. Neuronal Variability in Measures of Multisensory Integration

A single neuron is categorized as 'multisensory' if there is a statistically significant difference between the responses evoked by a multisensory stimulus combination and that evoked by the most effective of its individual components. The most widely known quantitative index expresses multisensory enhancement (or depression) as a proportion of the strongest unisensory response. Here we argue that, while having good descriptive value, this way of measuring interaction has several drawbacks. First, by only comparing mean responses to uni- *vs.* bimodal stimuli, it is not sensitive to changes in response variability. Second, it is lacking theoretical foundation in terms of the possible operations a neuron may perform in combining unisensory inputs to yield the multisensory response. Note that being responsive to several sensory modalities does not guarantee that a neuron has actually engaged in actively

integrating its multiple sensory inputs (Stanford *et al.*, 2005). As a result of these limitations, we suggest a new index grounded in a theoretical framework and that may be used to supplement more traditional measures. This new measure is as amenable to statistical testing as the more traditional measures and is completely analogous to a widely used procedure in behavioral studies of multisensory integration using the "race model inequality" (see Colonius and Diederich, 2006).

The new index proposed here is based on a probability summation mechanism. It compares the mean observed multisensory response of a neuron with the largest multisensory mean that is theoretically achievable by stochastically coupling its unisensory responses. Stochastic coupling means the construction of a joint probability distribution for two previously unrelated random variables (i.e., the unisensory response of the neuron to, say, a visual and an auditory stimulus). Due to coupling, one may examine the maximum of the two response random variables. Importantly, the maximum response will have the largest expected value (mean) when the visual and the auditory random variables have maximum negative stochastic dependency. This construction is known in statistics as the "method of antithetic variables" and is utilized in simulations to minimize variance (e.g., Ross, 2006). In practice, given a finite sample of responses to either modality, this construction is very simple. First, one orders both samples as column vectors (one column in increasing order, the second column in decreasing order) for the two unisensory components and then takes the maximum across the components (e.g., the maximum of largest visual response and the smallest auditory response, the second-largest visual response and the second-smallest auditory response, etc.). Finally, the mean of these maxima is computed, and this value represents the largest mean response that can be constructed using only the unisensory responses.

In analogy to the more traditional indices, the new index expresses the mean multisensory response as a proportion of the mean of the maximum of the unisensory responses constructed by the method of antithetic variables described above (see Fig. 1). It measures the degree by which a neuron's observed multisensory response surpasses the level obtainable by a theoretical neuron that optimally combines the unisensory responses under a probability summation rule (i.e., assuming that the neuron simply reacts to the more salient modality in any given multisensory trial). Importantly, and in contrast to the traditional indices, the new measure — being based on negative dependency coupling — capitalizes on the variability of the responses as measured (e.g., by their range): the larger the range in each of the unimodal responses, the larger the predicted mean of the maximum. In the extreme opposite case, with both unimodal responses being constant, the new index will not differ from the traditional one. Moreover, for the type of multisensory neurons that
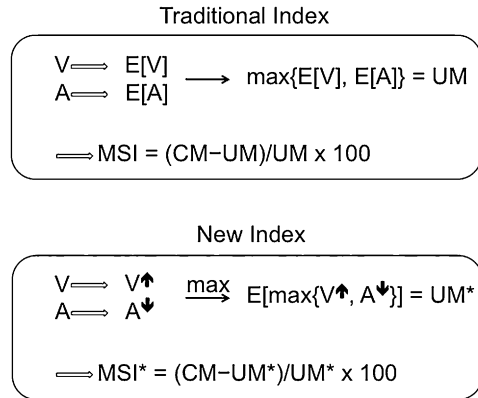
Traditional Index

New Index

Figure 1. Comparison of the traditional and a proposed new index of multisensory integration. The traditional index defines multisensory integration (MSI) as the difference between the mean of the actual cross modal response (CM) and the maximum mean unisensory response (UM). The UM is defined as the maximum of the expected values (means) from the individual visual (E[V]) and auditory response distributions (E[A]). The proposed new index (MSI*) first orders the individual trial by trial responses of the visual and auditory distributions in ascending and descending order, respectively. The maximum of each these pairings from the unisensory responses (largest visual response and smallest auditory response, secondly largest visual response and second smallest auditory response, etc.) is then taken, and the UM* is defined as the mean of these maxima, representing the largest mean response that could occur using unisensory inputs alone. Integration with the MSI* index is defined as the difference between UM* and the actual crossmodal mean response. Both measures are amenable to statistical testing.

do not respond at all to one of the single modalities, the two indexes will also take on the same value. The approach is illustrated by data from a single cat superior colliculus neuron (see Colonius and Diederich, 2015).

Because this new measure is, in general, more restrictive than the traditional one, many neurons categorized as 'multisensory' using more traditional measures may no longer fall into this category. However, it is similar to the upper bound of the race model inequality for reaction times under maximal negative dependence that is universally accepted as a valid test criterion, and therefore not an unreasonably high standard for 'multisensory'. Furthermore, as highlighted above, this measure provides a theoretical foundation for the possible operations a neuron may use to integrate unisensory inputs and produce a multisensory response, and thus could provide important mechanistic insights. Collectively, it is suggested that both this new and the more established measures of multisensory integration hold complementary value in providing important insights into both the cellular and network computations and the behavioral and perceptual benefits attributable to having information available from two or more senses.

3. Variability in Electroencephalography

The use of electroencephalographic recordings (EEG) has further illustrated the ubiquity and importance of within-subject variability in sensory perception (Mercier *et al.*, 2015; Morrell and Morrell, 1966; Sperdin *et al.*, 2009). Morrell and Morrell (1966) first provided evidence for differential neuronal processing at a within-subject level. These authors reported differences in terms of amplitude and peak latency of visual event-related potentials (ERPs) as a function of response times (RTs). Nonetheless, this study lacked the spatial resolution necessary to estimate the neuronal generators underlying such processing differences. Moreover, the authors only investigated differential behavioral and neuronal responses under visual-only stimulus presentations. In a more recent study, Sperdin and colleagues (2009), taking advantage of high-density EEG recordings, specifically investigated the non-linear multisensory interactions involved in fast *vs.* slow (median split of RTs) response trials. The results showed early supra-additive responses under multisensory presentations for fast response trials. Source estimations of these early effects located them to left temporal regions. Similarly, Mercier and colleagues (2015) recently provided evidence for supra-additive multisensory interactions within auditory cortices and that were differentially associated with fast *vs.* slow responses to auditory–visual stimuli. Although these studies were conducted within different subject cohorts (healthy volunteers *vs.* patient cohort), and differed in the choice of the dependent variable that was analyzed, there are intriguing similarities in the reported results. First, while Sperdin and colleagues (2009) investigated brain responses to auditory-somatosensory presentations, Mercier and colleagues (2015) report neuronal activity differences elicited by auditory–visual presentations. Interestingly, both studies highlight the distinct involvement of temporal cortices as a function of response speed. Secondly, Sperdin and colleagues (2009) reported effects collected at the global, scalp-evoked level (EEG), while Mercier and colleagues (2015) report data from intracranial recordings (electrocorticography (EcoG)). This difference in recording method could also lead to differences in the analysis strategy between these studies (global, topographic *vs.* local, oscillatory activity). Although the results were recorded at different neuronal scales and described by different dependent variables, the results can be reconciled within a common framework. Specifically, both studies focused on elucidating the neuronal networks that are selectively recruited under multisensory *vs.* unisensory presentation conditions (by only considering non-linear effects). However, less is know about the temporal dynamics underpinning within-subject RT variability under multisensory and unisensory presentation conditions.

The work presented during this symposium focused on directly identifying the neuronal networks differentially recruited on a trial-by-trial basis and that

result in substantial changes in RTs under both multisensory and visual-only presentation conditions. In addition, we sought to understand how individual differences in the magnitude of response speed variability affects these networks (Thelen, Ionta, Wallace and Murray, unpublished material). To this end, we re-analyzed a previously published ERP data set (Cappe *et al.*, 2010, 2012), this time specifically focusing on the neuronal networks that are differentially involved in RT.

More precisely, data from eight healthy individuals aged between 18 and 28 years (mean age 23 years) were considered in the analysis. The original experiment reported by Cappe and colleagues (Cappe *et al.*, 2009, 2012) comprised a total of 15 different stimulus conditions. In the present re-analysis of the data set, we only considered data from Go trials, where subjects reported the presence of a motion stimulus. Moreover, we did not include trials where motion stimuli had been presented in only one (auditory or visual) modality. This choice was necessary, in order to be able to compute race models for each multisensory condition. Thus, data from unisensory (auditory and visual receding or looming) and multisensory (auditory–visual, congruent and incongruent) were reanalyzed. In a first step, behavioral and EEG data were extracted and binned at a within-subject level, as function of the response time distributions. In other words, we separated the data of each subject according to whether the response time on a given trial fell within the first, second, third, or forth quartile of the response time distribution of a given condition (within-condition differences).

In a second step, we contrasted trials within the first quarter of the individual response time distributions with those in the last quarter of the response time distributions (i.e., first *vs.* last quartile). Analyses of the global electric field of the EEG signal revealed that RT variability co-varied with differential recruitment of an occipital-to-frontal network starting at ∼143 ms post-stimulus onset. More precisely, we found that response speed variability across subjects (i.e., leading to fast *vs.* slow response trials) was linked to differential recruitment of networks within visual, parietal, middle and inferior frontal cortices depending on the sensory presentation context (i.e., multisensory *vs.* visual-only). Under multisensory conditions, within-subject RT variability was linked to activation differences within middle frontal areas. However, RT variability under visual-only presentation conditions was linked to stronger recruitment of a right-hemispheric network involving occipito-temporal, parietal and frontal regions. This network differentiation suggests that within-subject RT variability under multisensory conditions seems to be linked to higher-level, top-down related processes, mediated by frontal cortices. Alternatively, RT variability under visual-only conditions seems to involve differential recruitment of both sensory bottom-up and higher-level top-down processes. A possible interpretation of these results is that multisensory presentations generally

enhance perception as indicted by the fact that there was no difference in the recruitment of lower-level cortices under fast *vs.* slow responses under auditory–visual presentations, converse to what was observed under visual-only conditions (Thelen *et al.*, unpublished material). Moreover, differences under multisensory conditions seem to be linked to higher-level, attentional and/or decisional processes as indicated by the stronger recruitment of middle frontal areas under slow responses to multisensory stimuli (see Yarkoni *et al.*, 2009 for a similar discussion).

Overall, we show that RT variability is linked to latency shifts (i.e., earlier for faster response times) in the recruitment of an extensive network spanning from early, low-level visual areas to pre-motor cortices. Such results are congruent with proposed models in which the degree of coordination of oscillatory activity within distant cortical areas acts as the underlying mechanisms mediating variability in information transfer across widespread neuronal networks (Engel *et al.*, 2001; Fries, 2015).

4. Variability in Functional Magnetic Resonance Imaging Data

A traditional strategy in neuroscience research is to look for overarching patterns of behavioral and neural responses to various classes of stimuli within various experimental paradigms. In functional magnetic resonance imaging (fMRI) studies in particular, a common approach is to transform individual subjects' data to a standard brain template (e.g., Talairach), which allows for group statistics to be calculated in a voxelwise manner. This assumes that evoked activity will occur in the same location in each person (i.e., that every brain is structurally and functionally identical). However, this assumption is often not only wrong but can lead to incorrect interpretations from analyzed data. For example, it is well known that the multisensory portion of the superior temporal sulcus (STS) is in a slightly different location within the anatomical STS in individual subjects, which might allow for an effect to be washed out when individual maps are averaged in standard space. Although the whole-brain analysis approach is useful for understanding general patterns and concepts in neuroscience, it frequently overlooks that across any set of trials or subjects there is marked variability. We can ask the question, how much variability exists within and across individuals in a multisensory context? Furthermore, how are our findings affected by analytical choices like performing our analyses in standard space as opposed to individual subjects' native image space?

To investigate these questions, we utilized a traditional redundant target paradigm in which the intensity of the auditory (1000 Hz beep) and visual (white flash against a gray background) stimuli were parametrically manipulated to create a three by three (no stimulus, faint stimulus, suprathreshold

stimulus along both auditory and visual modalities) design yielding auditory alone, visual alone, and audiovisual conditions. Subjects ($n = 14$, nine female, mean age 25.5 years) were asked to detect the target stimulus as quickly and accurately as possible, and completed the task while in the MRI environment. Each run lasted approximately four minutes and five scan runs were collected from each subject. In each run, targets were presented in a rapid-event related design and included 25% catch trials. Data were collected from a 3T Philips Achieva scanner using a 32-channel head coil. In addition to the functional data, two T1-weighted MP-RAGE anatomical scans were collected from each subject. All data were analyzed using Analysis of Functional NeuroImages (AFNI) (Cox, 1996).

Behavioral results revealed the expected pattern of faster reaction times as both auditory and visual stimulus intensity was increased. A two-way ANOVA with auditory level as one factor and visual level as a second factor revealed a main effect of both auditory and visual level and an interaction between factors for both detection and reaction time data. However, BOLD responses revealed several interesting response patterns. Using a whole-brain analysis, key parts of the robust and well-known audiovisual multisensory perceptual network were not significantly activated in the task. Although there was significant sensory evoked activity across the brain, in a whole-brain analysis using 3dANOVA2 in AFNI, there were no areas in primary auditory cortex that showed a main effect of auditory level (volume) in the BOLD response. Similarly, there were no areas in visual cortex that showed a main effect of visual level (brightness) in the BOLD response. However, when a region of interest (ROI) analysis was conducted with ROIs identified in individual subjects in native image space, we found a much more expected pattern of results. For example, both right and left primary auditory cortex showed a main effect of auditory level, revealing a robust and expected effect that the whole-brain analysis failed to show.

We next turned our attention to the relationship between the neural response and behavioral performance in individual subjects. On the group level, there was no clear relationship between activation pattern in any ROI and response time. In individual subjects, however, we found evidence for a neural-behavioral relationship that varied across subjects. For example, one subject showed a linear relationship between response time and percent signal change in the left superior temporal sulcus (STS), with faster reaction times associated with decreased response amplitude. Another subject showed the opposite pattern, with a linear relationship showing slower reaction times associated with increased response amplitude. Further work is needed to ascribe a mechanism for this difference. For example, individual differences from the group average may represent spurious fluctuations that might fall within the range of variability expected between individual scans of the same participant. To

determine the stability of individual differences in response patterns, a representational similarity analysis (RSA) can be done to tease apart spurious *vs.* consistent individual differences (Kriegeskorte *et al.*, 2008).

Our work shows the need to not only examine behavioral and fMRI data on a group level with a 'big picture' view of behavioral and neural response patterns, but rather to also examine responses on an individual subject level. This may reveal interesting subsets of subjects that show distinct patterns of activation associated with a behavioral task.

5. Variability in Electrocorticography Responses

Electrocorticography (ECoG) in awake human patients represents a powerful method to examine information encoding with good spatial and excellent temporal resolution, and can also be valuable in assessing the role of variability in neural and behavioral responses. In particular, this human electrophysiology technique allows the recording of neuronal activity in near proximity to the current generators. In fact the recording leads are laid over the brain's pial surface, which was previously exposed for epilepsy monitoring. It is therefore apparent that the signal-to-noise ratio of ECoG recordings is spatially more specific and its magnitude is higher as opposed to other methods of human neuroimaging (e.g., EEG, fMRI). We asked how the variability of neuronal activity across different areas and in different subjects (spatial specificity) could help us in the interpretation of multisensory stimuli integration in the brain.

In an ECoG study, we sought to investigate whether neural representations of auditory speech features in posterior superior temporal cortex (pSTG) are altered by concurrent visual speech, and the contributions of variability to these processes. We presented speech in two experimental conditions to seven patients implanted with subdural electrodes for seizure monitoring. In the auditory-only (A) condition, videos of complete spoken sentences were presented with a static face of the speaker. In the audiovisual (AV) condition, videos of complete sentences were presented with matching dynamic visual speech information. After each sentence, a single word was presented and the subject had to indicate whether the word had occurred in the previous sentence or not. This task ensured that the subject paid attention to each word in the sentence.

To determine the effect of visual speech on auditory speech responses in pSTG, two different analyses were employed that probed for effects at different levels of representational complexity. First, we compared amplitudes of neural responses elicited by auditory-only and audiovisual speech conditions in electrodes covering pSTG at the single subject level. To this aim we used a standard pipeline according to FieldTrip (Oostenveld *et al.*, 2011), a

Matlab toolbox for time series analysis (parameters: independent samples *t*-test on magnitude differences, alpha = 0.05 uncorrected, magnitude frequency bands from Fourier analysis: 50–170 Hz in steps of 2 Hz, time of interest: 300–2500 ms after speech onset, in steps of 50 ms, window length 250 ms, multitapering with four tapers). Second, we correlated the speech envelope with the band-passed time courses of the single subject's ECoG signal on a trial-by-trial basis to detect which electrodes followed variations of acoustic energy in time (time of interest: 300–2500 ms in steps of 10 ms and window length 100 ms, with each frequency bin's time course is forward/backward delayed in steps of 10 ms and correlated with the speech envelope).

Across all subjects we found neural activity changes in pSTG electrodes. For the first analysis, we found a gamma power (50–70 Hz) increase in the response to both A and AV speech in the pSTG (300–2500 ms post stimulus onset) relative to baseline, with a larger power increase for AV speech than A speech (86% *vs.* 78%, all *p* values < 0.05, unpaired *t*-test within each of five electrodes). In comparison, anterior STG electrodes did not show significant difference between AV and A conditions.

For the second analysis, we found that for both AV and A speech, high gamma power (70–170 Hz) in pSTG electrodes tracked the auditory speech envelope. Using a simple analysis, we found no difference in the highest correlation with the speech envelope between the AV and the A condition (max $r = 0.75 \pm 0.01$ SEM for AV and 0.76 ± 0.01 SEM for A). The electrode that showed the highest correlation with the speech envelope was the same that showed the largest power changes (gamma and beta bands) in either AV or A condition. Behavioral performance was high in both conditions (87–90% correct performance) showing that subjects attended to the stimulus.

Visual speech provides an additional independent source of information about the content of speech. pSTG may be a neural locus/hub of multisensory speech perception (Nath and Beauchamp, 2011). In accordance, we found differential responses (greater gamma activity) for AV than A speech in pSTG. In the second analysis, we found that gamma activity tracked the speech envelope in agreement with previous observations (Kubanek *et al.*, 2013), however, we did not find evidence for enhanced envelope tracking for AV compared with A speech. This could be because behavioral performance was very high in both conditions. Because noisy auditory speech shows the greatest enhancement for visual speech (Ross *et al.*, 2007), the next experiments will compare envelope tracking for clear and noisy audiovisual speech.

Our project shows that the variability of electrodes coverage in ECoG can be a limiting factor in the detection of multisensory integration effects. Conversely, the advantage of such a high signal-to-noise ratio allows us to exploit the variability of different oscillatory rhythms, such as in beta and gamma bands, and is advantageous in the encoding of information related to stimuli

features such as speech envelope. In fact, the variable power activation of beta and gamma bands across different electrodes, might reveal speech envelope encoding in either low (5–30 Hz) or gamma, or both bands together.

In conclusion, with the current methodology we provide evidence of audiovisual interactions in pSTG (five or six electrodes, two or three subjects, two gamma bands) that may not be due to the presence of the additional visual speech information alone but rather an integration of the auditory and visual speech information. The audiovisual effect is supported by the presence of speech envelope correlation in the mentioned pSTG electrodes, but lack thereof in the visual electrodes. Future work will look at the phase of neural responses to determine whether visual speech alters the phase in higher order auditory cortex and we will evaluate how well the brain encodes speech features *via* receptive fields' estimation.

6. Discussion

As evidenced by the work described in the above sections, both inter- and intrasubject variability have been underappreciated in the context of multisensory function, and therefore a greater understanding of variability holds the potential to provide important mechanistic insights into multisensory encoding and function. At each scale of analysis discussed here, which ranges from single neurons to whole brain, it is clear that substantial variance is the rule rather than the exception. Avenues of future work will be to detail how variance at these different spatial and spatiotemporal scales contributes to differences in behavioral and perceptual performance (illustrated in several of the examples presented here), and how variability at each level contributes to variability at the next level.

The variability in response that is seen within subjects at the trial-by-trial level is the likely product of a host of interacting factors. Thinking first from the viewpoint of a single neuron, differences in spiking in response to an identical stimulus are likely due to differences in the membrane potential of the neuron and/or differences in the synaptic drive coming from inputs onto that neuron. These in turn could be the result of differences in factors such as noise, phase within an oscillatory cycle, state, prior history, etc. When scaled to the level of local ensemble encoding and larger networks, additional sources of variance emerge in factors such as strength of oscillatory coupling, network dynamics and slight connectivity differences. Indeed, when thinking about the vast array of non-stimulus dependent features that are at play at any given time in a neuron or neural network, it is easy to understand significant variability on a trial-by-trial basis.

Although interindividual variability in sensory and multisensory responsiveness is a natural and emergent characteristic of complex biological systems, there are likely limits to this variability in relation to ideal encoding. On the one hand, 'normal' variability is beneficial by providing some degree of flexibility in encoding, and thus may represent the basis for adaptive plasticity. On the other hand, excessive variability is likely to compromise the encoding processes, thus creating 'fuzzy' perceptual representations that are poor maps of the (multi)sensory world. Emerging evidence suggests that excessive variability may be an important contributing factor in cognitive differences ranging from autism to pathological aging. Future work is necessary to understand the contributions of variability within multisensory processes and networks to the perceptual and cognitive deficits that characterize these disorders.

Finally, much work still needs to be done and that focuses on linking trial-by-trial fluctuations in neuronal responses to trial-by-trial differences in behavior and perception. Such experiments and their associated analyses hold great promise in elucidating the contributions of given brain regions and networks to specific behavioral, perceptual and cognitive facilities. Indeed, experiments can be structured such as to take advantage of the inherent variability in brain responses by, for example, examining the differences in activation patterns in response to identical multisensory stimulus combinations that result in different behaviors and percepts.

References

Adamo, N., Huo, L., Adelsberg, S., Petkova, E., Castellanos, F. X. and Di Martino, A. (2014). Response time intra-subject variability: commonalities between children with autism spectrum disorders and children with ADHD, *Eur. Child Adolesc. Psychiatry* **23**, 69–79.

Alais, D., Newell, F. N. and Mamassian, P. (2010). Multisensory processing in review: from physiology to behaviour, *Seeing Perceiving* **23**, 3–38.

Baum, S. H. and Beauchamp, M. S. (2014). Greater BOLD variability in older compared with younger adults during audiovisual speech perception, *PLoS ONE* **9**, e111121. DOI:10.1371/journal.pone.0111121.

Baum, S. H., Stevenson, R. A. and Wallace, M. T. (2015). Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder, *Progr. Neurobiol.* **134**, 140–160.

Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E. and Pouget, A. (2012). Not noisy, just wrong: the role of suboptimal inference in behavioral variability, *Neuron* **74**, 30–39.

Bielak, A. A., Hultsch, D. F., Strauss, E. H., Macdonald, S. W. and Hunter, M. A. (2010a). Intraindividual variability is related to cognitive change in older adults: evidence for within-person coupling, *Psychol. Aging* **25**, 575–586.

Bielak, A. A., Hultsch, D. F., Strauss, E. H., Macdonald, S. W. and Hunter, M. A. (2010b). Intraindividual variability in reaction time predicts cognitive outcomes 5 years later, *Neuropsychology* **24**, 731–741.

Calvert, G. A., Spence, C. and Stein, B. E. (2004). *Handbook of Multisensory Processes*. MIT Press, Cambridge, MS, USA.

Cappe, C., Thut, G., Romei, V. and Murray, M. M. (2009). Selective integration of auditory–visual looming cues by humans, *Neuropsychologia* **47**, 1045–1052.

Cappe, C., Thut, G., Romei, V. and Murray, M. M. (2010). Auditory–visual multisensory interactions in humans: timing, topography, directionality, and sources, *J. Neurosci.* **30**, 12572–12580.

Cappe, C., Thelen, A., Romei, V., Thut, G. and Murray, M. M. (2012). Looming signals reveal synergistic principles of multisensory integration, *J. Neurosci.* **32**, 1171–1182.

Cerella, J. and Hale, S. (1994). The rise and fall in information-processing rates over the life span, *Acta Psychol.* **86**, 109–197.

Christen, M., Kohn, A., Ott, T. and Stoop, R. (2006). Measuring spike pattern reliability with the Lempel–Ziv-distance, *J. Neurosci. Methods* **156**, 342–350.

Christensen, K., Olami, Z. and Bak, P. (1992). Deterministic $1/f$ noise in nonconserative models of self-organized criticality, *Phys. Rev. Lett.* **68**, 2417–2420.

Colonius, H. and Diederich, A. (2006). The race model inequality: interpreting a geometric measure of the amount of violation, *Psychol. Rev.* **113**, 148–154.

Colonius, H. and Diederich, A. (2015). A new measure of multisensory integration in a single neuron based on dependent probability summation. Available at arXiv:1507.08505v1.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages, *Comput. Biomed. Res.* **29**, 162–173.

Di Martino, A., Ghaffari, M., Curchack, J., Reiss, P., Hyde, C., Vannucci, M., Petkova, E., Klein, D. F. and Castellanos, F. X. (2008). Decomposing intra-subject variability in children with attention-deficit/hyperactivity disorder, *Biol. Psychiatry* **64**, 607–614.

Dutta, P. and Horn, P. M. (1981). Low-frequency fluctuations in solids: $1/f$ noise, *Rev. Modern Phys.* **53**, 497–499.

Engel, A. K., Fries, P. and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top–down processing, *Nat. Rev. Neurosci.* **2**, 704–716.

Fries, P. (2015). Rhythms for cognition: communication through coherence, *Neuron* **88**, 220–235.

Garrett, D. D., Kovacevic, N., Mcintosh, A. R. and Grady, C. L. (2011). The importance of being variable, *J. Neurosci.* **31**, 4496–4503.

Garrett, D. D., Macdonald, S. W. and Craik, F. I. (2012). Intraindividual reaction time variability is malleable: feedback- and education-related reductions in variability with age, *Front. Hum. Neurosci.* **6**, 101. DOI:10.3389/fnhum.2012.00101.

Hedden, T. and Gabrieli, J. D. (2004). Insights into the ageing mind: a view from cognitive neuroscience, *Nat. Rev. Neurosci.* **5**, 87–96.

Hultsch, D. F., Macdonald, S. W. and Dixon, R. A. (2002). Variability in reaction time performance of younger and older adults, *J. Gerontol. B, Psychol. Sci. Soc. Sci.* **57**, P101–P115.

Jenkins, L., Myerson, J., Joerding, J. A. and Hale, S. (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition, *Psychol. Aging* **15**, 157–175.

Kriegeskorte, N., Mur, M. and Bandettini, P. (2008). Representational similarity analysis — connecting the branches of systems neuroscience, *Front. Syst. Neurosci.* **2**, 4. DOI:10.3389/neuro.06.004.2008.

Kubanek, J., Brunner, P., Gunduz, A., Poeppel, D. and Schalk, G. (2013). The tracking of speech envelope in the human cortex, *PLoS ONE* **8**, e53398. DOI:10.1371/journal.pone.0053398.

Lovden, M., Li, S. C., Shing, Y. L. and Lindenberger, U. (2007). Within-person trial-to-trial variability precedes and predicts cognitive decline in old and very old age: longitudinal data from the Berlin Aging Study, *Neuropsychologia* **45**, 2827–2838.

Lovden, M., Schmiedek, F., Kennedy, K. M., Rodrigue, K. M., Lindenberger, U. and Raz, N. (2013). Does variability in cognitive performance correlate with frontal brain volume? *NeuroImage* **64**, 209–215.

MacDonald, S. W., Karlsson, S., Rieckmann, A., Nyberg, L. and Backman, L. (2012). Aging-related increases in behavioral variability: relations to losses of dopamine D1 receptors, *J. Neurosci.* **32**, 8186–8191.

Mercier, M. R., Molholm, S., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H. and Foxe, J. J. (2015). Neuro-oscillatory phase alignment drives speeded multisensory response times: an electro-corticographic investigation, *J. Neurosci.* **35**, 8546–8557.

Morrell, L. K. and Morrell, F. (1966). Evoked potentials and reaction times: a study of intra-individual variability, *Electroencephalogr. Clin. Neurophysiol.* **20**, 567–575.

Murphy, K. J., West, R., Armilio, M. L., Craik, F. I. and Stuss, D. T. (2007). Word-list-learning performance in younger and older adults: intra-individual performance variability and false memory, *Neuropsychol., Dev. Cogn. B Aging Neuropsychol. Cogn.* **14**, 70–94.

Murray, M. M., Foxe, J. J. and Wylie, G. R. (2005). The brain uses single-trial multisensory memories to discriminate without awareness, *NeuroImage* **27**, 473–478.

Murray, M. M. and Wallace, M. T. (2012). *The Neural Bases of Multisensory Processes*. CRC Press, Boca Raton, FL, USA.

Nath, A. R. and Beauchamp, M. S. (2011). Dynamic changes in superior temporal sulcus connectivity during perception of noisy audiovisual speech, *J. Neurosci.* **31**, 1704–1714.

Ohtsuki, H., Hauert, C., Lieberman, E. and Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks, *Nature* **441**, 502–505.

Oostenveld, R., Fries, P., Maris, E. and Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data, *Comput. Intell. Neurosci.* **2011**, 156869. DOI:10.1155/2011/156869.

Peich, M. C., Husain, M. and Bays, P. M. (2013). Age-related decline of precision and binding in visual working memory, *Psychol. Aging* **28**, 729–743.

Romei, V., Murray, M. M., Cappe, C. and Thut, G. (2013). The contributions of sensory dominance and attentional bias to cross-modal enhancement of visual cortex excitability, *J. Cogn. Neurosci.* **25**, 1122–1135.

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C. and Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments, *Cereb. Cortex* **17**, 1147–1153.

Ross, S. M. (2006). *Simulation*. Academic Press, Burlington, MA, USA.

Sarko, D. K., Nidiffer, A. R., Powers, I. A., Ghose, D., Hillock-Dunn, A., Fister, M. C., Krueger, J. and Wallace, M. T. (2012). Spatial and temporal features of multisensory processes: bridging animal and human studies, in: *The Neural Bases of Multisensory Processes*, M. M. Murray and M. T. Wallace (Eds), pp. 192–216. CRC Press, Boca Raton, FL, USA.

Sperdin, H. F., Cappe, C., Foxe, J. J. and Murray, M. M. (2009). Early, low-level auditory-somatosensory multisensory interactions impact reaction time speed, *Front. Integr. Neurosci.* **3**, 2. DOI:10.3389/neuro.07.002.2009.

Stanford, T. R., Quessy, S. and Stein, B. E. (2005). Evaluating the operations underlying multisensory integration in the cat superior colliculus, *J. Neurosci.* **25**, 6499–6508.

Stein, B. E. and Meredith, M. A. (1993). *The Merging of the Senses*. MIT Press, Cambridge, MA, USA.

Stevenson, R. A. and Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies, *Exp. Brain Res.* **227**, 249–261.

Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., Kurela, L. R., Siemann, J. K., James, T. W. and Wallace, M. T. (2014). Identifying and quantifying multisensory integration: a tutorial review, *Brain Topogr.* **27**, 707–730.

Thelen, A., Matusz, P. J. and Murray, M. M. (2014). Multisensory context portends object memory, *Curr. Biol.* **24**, R734–R735.

Voss, R. F. (1992). Evolution of long-range fractal correlations and $1/f$ noise in DNA base sequences, *Phys. Rev. Lett.* **68**, 3805–3808.

Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W. and Schirillo, J. A. (2004). Unifying multisensory signals across time and space, *Exp. Brain Res.* **158**, 252–258.

West, R., Murphy, K. J., Armilio, M. L., Craik, F. I. and Stuss, D. T. (2002). Lapses of intention and performance variability reveal age-related increases in fluctuations of executive control, *Brain Cogn.* **49**, 402–419.

Yarkoni, T., Barch, D. M., Gray, J. R., Conturo, T. E. and Braver, T. S. (2009). BOLD correlates of trial-by-trial reaction time variability in gray matter and white matter: a multi-study fMRI analysis, *PLoS ONE* **4**, e4257. DOI:10.1371/journal.pone.0004257.