The role of prediction and attention in phantom voice perception

1. Summary

Modulations of the interplay of sensory prediction and attentional control are often associated with the experience of hearing voices (Ford, Gray, Faustman, Roach, & Mathalon, 2007; Griffin & Fletcher, 2017; Heinks-Maldonado et al., 2007; Hugdahl et al., 2008; Nelson, Whitford, Lavoie, & Sass, 2014a, 2014b). Unlike non-clinical voice hearers, voice hearers with a diagnosis of psychotic disorder more often perceive derogatory voices, thus further accentuating an individual's attentional bias towards negative emotions (Daalman et al., 2011; Johns et al., 2014; Nelson et al., 2014a). This negative attentional bias might also lead to imprecise or aberrant predictions of negative cues and result in misattribution of negative meaning (threat) to a neutral stimulus (Alba-Ferrara, de Erausquin, Hirnstein, Weis, & Hausmann, 2013; Nelson et al., 2014a). Similarly, imprecise or aberrant predictive processing might contribute to the inability to differentiate internally- from externally-generated events (Griffin & Fletcher, 2017; Nelson et al., 2014b). However, it is still unclear how modulations of predictive processing and attentional control might influence the transition from non-clinical to clinically relevant AVH. Chapter 2 elucidated if and how negative and positive emotional stimuli control attention allocation to task-relevant and task-irrelevant aspects in a conflict scenario. Specifically, it explored whether distinct behavioral and neural responses arise from these emotions or if they simply mark salience. An adapted version of the flanker task was combined with fMRI to investigate valence-specific emotion effects on attentional control in conflict processing. Slower behavioral responses were observed for high (incongruent) compared to low (congruent) conflict scenarios. However, negative and positive emotions did not evoke distinct responses. Neural activity in the dorsal ACC pointed toward its general role in monitoring and assessing conflict as well as in selecting appropriate responses regardless of the stimulus quality. Findings from this study thus confirmed that negative and positive emotional stimuli mark salience in both low (congruent) and high (incongruent) conflict scenarios. This suggests that, regardless of the conflict level, emotional stimuli attract more attentional resources in goal-directed behavior than neutral stimuli. The lack of significant differences between negative and positive emotional stimuli could be attributed to switching between trials of different congruence, arousal, and valence, which might have created an experimental context that required higher cognitive effort to sustain attentional control. These switching costs also

might have diluted valence-specific effects on conflict processing. Overall, the findings underscore the potential of attentional control to mitigate the impact of emotional contexts in both high and low conflict situations, thereby facilitating the attainment of overall task goals.

Behavioral and neural changes in emotion processing and attentional control are often described in non-clinical voice hearers and those with a diagnosis of a psychotic disorder (Alba-Ferrara et al., 2013; Amminger et al., 2012; Amorim, Roberto, Kotz, & Pinheiro, 2022; Pinheiro & Niznikiewicz, 2019; Pinheiro et al., 2014; Pinheiro et al., 2017). Chapter 3 discussed and integrated fMRI evidence pertaining to the attentional control of emotion along a postulated psychosis continuum. Over-sensitivity towards attentional control of negative compared to positive emotions was associated with the severity of positive and negative symptoms in voice hearers. Further, the interaction between emotion and context-sensitive attentional control was altered and differentially influenced by factors such as arousal, motivation, and reward. This might contribute to altered interpersonal communication and real-life skills in individuals with psychosis. Lastly, increased neural activity in subcortical brain regions that mediate the coupling of emotion and cognitive control (e.g., basal ganglia, thalamus, and angular gyrus) was associated with increased effort in emotion-attentional control tasks in psychosis, reflecting the role and influence of context and individual differences on the interplay of emotion processing and attentional control (chapter 3). However, further research is necessary to specify how these findings might contribute to the understanding of transitions from non-clinical to clinically relevant positive symptoms.

To acquire a more thorough understanding of the contributions of attentional control, emotion processing, and sensory predictive processing in voice hearing, the emotional quality of the self-voice was systematically altered (chapter 4, 6 – neutral-angry; chapter 5 – neutral-pleasure) to manipulate the degree of certainty about sensory feedback in self-voice production. ERPs evoked by an individual's self- and externally-generated self-voice (unambiguous/certain [100% neutral, 100% emotional] and ambiguous/uncertain [60-40% neutral- emotional, 50-50% neutral- emotional, 40-60% neutral- emotional]) were examined as a function of HP (based on LSHS scores) in a classical auditory-motor button-press paradigm. The N100 (chapter 4, 6) and P200 (chapter 6) suppression effects were replicated such that their amplitudes were suppressed for self- compared to externally-generated self-voice, regardless of their emotional quality or the degree of HP.

The N100 suppression effect is related to reduced activity in the auditory cortex when expected sensations match the actual sensory feedback of one's own voice compared to when it is passively listened to (Ford & Mathalon, 2004; Ford et al., 2001; Ford et al., 2013; Pinheiro, Schwartze, & Kotz, 2018). Similarly, the P200 suppression likely reflects the conscious distinction between the self- and externally-generated self-voice (Knolle, Schroger, Baess, & Kotz, 2012; Sowman, Kuusik, & Johnson, 2012). Conversely, increased N100/P200 responses to the externally- compared to self-generated voice can also be interpreted as an increase in attention allocation to the unpredictable onset of the self-voice. The P50 response (chapter 6) was higher for self- compared to externally-generated self-voices, likely indicating higher levels of alertness associated with anticipation of the novelty and unexpectedness of different voices/voice feedback presented in the self-generated condition, which preceded the externally-generated condition in the blocked task design (Bramon, Rabe-Hesketh, Sham, Murray, & Frangou, 2004; Grimm & Escera, 2012; Patterson et al., 2008; Thaker, 2008). To examine whether changes in the certainty of sensory feedback to self-voices recruit additional attentional resources, the N100 and N200 responses of both self-generated and externally-generated voices were correlated. The N100 and the N200 responses for self-generated certain/unambiguous but not for uncertain/ambiguous self-voice correlated significantly (see chapter 6: figure 7). This significant correlation might reflect engagement of additional resource allocation for self-relevant unambiguous self-voices (100% neutral and 100% angry self-voice), whereas the lack of significant correlation might reflect disengagement of resource allocation for further perceptual processing of non-self-relevant ambiguous voices (60-40% neutral-angry and 50-50% neutral-angry self-voice). HP modulated the P50 (chapter 6), N100 (chapter 4, 6) and P200 (chapter 6) suppression effects such that the N100 suppression effect decreased (MAc > AO), whereas the P50 and P200 suppression effects increased (AO > MAc) with higher HP scores. The decreased or reversed N100 suppression effect, i.e., an increased response to the self-generated voice in high HP individuals, reflects altered sensory feedback processing and suggests a discrepancy in the matching of the expected and actual sensory feedback of the self-voice (Ford & Mathalon, 2004; Ford et al., 2001; Ford et al., 2013; Pinheiro et al., 2018). Alternatively, it may indicate increased error awareness and/or attention allocation to the self-generated own voice in high HP individuals. The larger P50 suppression effect, specifically increased responses during the externally-generated condition in high HP, may reflect larger effort to remain alert and process an overload of sensory information, i.e., the variation in emotional voice quality as well as in the recognition of the self-voice (Bramon et al., 2004; Grimm & Escera, 2012; Patterson et al., 2008; Thaker, 2008). Similarly, increased P200 responses to externally-generated self-voices in high HP may also indicate increased attentional allocation and more conscious processing of the self-voice differing in emotional quality. Of note is that the effects of changes in HP on sensory feedback processing and attentional control were only observed in the context of the neutral-angry emotion spectrum in self-voice production (chapter 4, 6), but not in the neutral-pleasure emotional context (chapter 5). This might be due to lower perceptual discriminability among the five types of voices varying in pleasure content (i.e., similar acoustic properties), which may have resulted in an ambiguous context. This low perceptual discriminability and ambiguous context may have resulted in the lack of discernible differences in certainty about the sensory feedback to self-voices and consequently, a lack of differences in attentional engagement.

To conclude, the combined findings suggest that (i) efficient attentional control counteracts the distracting influence of emotional stimuli on attaining task goals (chapter 2); (ii) alterations in the context-sensitive attentional control of emotion are associated with increased HP (chapter 3); (iii) increased HP is associated with (neurophysiological) changes in sensory prediction and feedback processing as well as attentional control in self-voice production (chapter 4, 6). These findings complement the psychosis continuum hypothesis, identifying alterations in the processing of sensory consequences of one's own actions as well as attentional control in individuals with high HP, which were previously reported in individuals with psychotic disorder (chapter 4, 6). Overall, these findings critically advance the understanding of the neural dynamics of voice hearing and AVH.

REFERENCES

Alba-Ferrara, L., de Erausquin, G. A., Hirnstein, M., Weis, S., & Hausmann, M. (2013). Emotional prosody modulates attention in schizophrenia patients with hallucinations. *Front Hum Neurosci, 7*, 59. doi: 10.3389/fnhum.2013.00059

Amminger, G. P., Schafer, M. R., Papageorgiou, K., Klier, C. M., Schlogelhofer, M., Mossaheb, N., . . . McGorry, P. D. (2012). Emotion recognition in individuals at clinical high-risk for schizophrenia. *Schizophr Bull, 38*(5), 1030-1039. doi: 10.1093/schbul/sbr015

Amorim, M., Roberto, M. S., Kotz, S. A., & Pinheiro, A. P. (2022). The perceived salience of vocal emotions is dampened in non-clinical auditory verbal hallucinations. *Cogn Neuropsychiatry, 27*(2-3), 169-182. doi: 10.1080/13546805.2021.1949972

Bramon, E., Rabe-Hesketh, S., Sham, P., Murray, R. M., & Frangou, S. (2004). Meta-analysis of the P300 and P50 waveforms in schizophrenia. *Schizophr Res, 70*(2-3), 315-329.

Daalman, K., Boks, M. P., Diederen, K. M., de Weijer, A. D., Blom, J. D., Kahn, R. S., & Sommer, I. E. (2011). The same or different? A phenomenological comparison of auditory verbal hallucinations in healthy and psychotic individuals. *J Clin Psychiatry, 72*(3), 320-325. doi: 10.4088/JCP.09m05797yel

Ford, J. M., Gray, M., Faustman, W. O., Roach, B. J., & Mathalon, D. H. (2007). Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology, 44*(4), 522-529. doi: 10.1111/j.1469-8986.2007.00533.x

Ford, J. M., & Mathalon, D. H. (2004). Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *J Psychiatr Res, 38*(1), 37-46. doi: 10.1016/s0022-3956(03)00095-5

Ford, J. M., Mathalon, D. H., Kalba, S., Whitfield, S., Faustman, W. O., & Roth, W. T. (2001). Cortical responsiveness during talking and listening in schizophrenia: an event-related brain potential study. *Biol Psychiatry, 50*(7), 540-549. doi: 10.1016/s0006-3223(01)01166-0

Ford, J. M., Mathalon, D. H., Roach, B. J., Keedy, S. K., Reilly, J. L., Gershon, E. S., & Sweeney, J. A. (2013). Neurophysiological evidence of corollary discharge function during vocalization in psychotic patients and their nonpsychotic first-degree relatives. *Schizophr Bull, 39*(6), 1272-1280. doi: 10.1093/schbul/sbs129

Griffin, J. D., & Fletcher, P. C. (2017). Predictive Processing, Source Monitoring, and Psychosis. *Annu Rev Clin Psychol, 13*, 265-289. doi: 10.1146/annurev-clinpsy-032816-045145

Grimm, S., & Escera, C. (2012). Auditory deviance detection revisited: evidence for a hierarchical novelty system. *International Journal of Psychophysiology, 85*(1), 88-92.

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., & Ford, J. M. (2007). Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Arch Gen Psychiatry, 64*(3), 286-296. doi: 10.1001/archpsyc.64.3.286

Hugdahl, K., Løberg, E.-M., Specht, K., Steen, V. M., Wageningen, H. v., & Jørgensen, H. A. (2008). Auditory hallucinations in schizophrenia: the role of cognitive, brain structural and genetic disturbances in the left temporal lobe. *Front Hum Neurosci, 2*, 131.

Johns, L. C., Kompus, K., Connell, M., Humpston, C., Lincoln, T. M., Longden, E., . . . Laroi, F. (2014). Auditory verbal hallucinations in persons with and without a need for care. *Schizophr Bull, 40 Suppl 4*(Suppl 4), S255-264. doi: 10.1093/schbul/sbu005

Knolle, F., Schroger, E., Baess, P., & Kotz, S. A. (2012). The cerebellum generates motor-to-auditory predictions: ERP lesion evidence. *J Cogn Neurosci, 24*(3), 698-706. doi: 10.1162/jocn_a_00167

Nelson, B., Whitford, T. J., Lavoie, S., & Sass, L. A. (2014a). What are the neurocognitive correlates of basic self-disturbance in schizophrenia? Integrating phenomenology and neurocognition: Part 2 (aberrant salience). *Schizophr Res, 152*(1), 20-27. doi: 10.1016/j.schres.2013.06.033

Nelson, B., Whitford, T. J., Lavoie, S., & Sass, L. A. (2014b). What are the neurocognitive correlates of basic self-disturbance in schizophrenia?: Integrating phenomenology and neurocognition. Part 1 (Source monitoring deficits). *Schizophr Res, 152*(1), 12-19. doi: 10.1016/j.schres.2013.06.022

Patterson, J. V., Hetrick, W. P., Boutros, N. N., Jin, Y., Sandman, C., Stern, H., . . . Bunney, W. E., Jr. (2008). P50 sensory gating ratios in schizophrenics and controls: a review and data analysis. *Psychiatry Res, 158*(2), 226-247. doi: 10.1016/j.psychres.2007.02.009

Pinheiro, A. P., & Niznikiewicz, M. (2019). Altered attentional processing of happy prosody in schizophrenia. *Schizophr Res, 206*, 217-224. doi: 10.1016/j.schres.2018.11.024

Pinheiro, A. P., Rezaii, N., Rauber, A., Liu, T., Nestor, P. G., McCarley, R. W., . . . Niznikiewicz, M. A. (2014). Abnormalities in the processing of emotional prosody from single words in schizophrenia. *Schizophr Res, 152*(1), 235-241. doi: 10.1016/j.schres.2013.10.042

Pinheiro, A. P., Rezaii, N., Rauber, A., Nestor, P. G., Spencer, K. M., & Niznikiewicz, M. (2017). Emotional self-other voice processing in schizophrenia and its relationship with hallucinations: ERP evidence. *Psychophysiology, 54*(9), 1252-1265. doi: 10.1111/psyp.12880

Pinheiro, A. P., Schwartze, M., & Kotz, S. A. (2018). Voice-selective prediction alterations in nonclinical voice hearers. *Sci Rep, 8*(1), 14717. doi: 10.1038/s41598-018-32614-9

Sowman, P. F., Kuusik, A., & Johnson, B. W. (2012). Self-initiation and temporal cueing of monaural tones reduce the auditory N1 and P2. *Exp Brain Res, 222*, 149-157.

Thaker, G. K. (2008). Neurophysiological endophenotypes across bipolar and schizophrenia psychosis. *Schizophr Bull, 34*(4), 760-773. doi: 10.1093/schbul/sbn049