**Comodulação da variabilidade da frequência cardíaca e linguagem**

**André Fujita**

*Universidade de São Paulo*

**Mayara dos Santos Nascimento**

*Universidade de São Paulo*

1. **Introdução**

Os seres humanos têm a capacidade única de comunicar e extrair significado através da linguagem falada e escrita. E tal processamento da linguagem é uma das tarefas cognitivas mais complexas em que os humanos se envolvem rotineiramente. Uma parte integrante da linguagem humana é a capacidade de extrair significado das palavras faladas e escritas, mas a relação precisa entre as representações cerebrais das informações percebidas pela audição versus leitura não é clara.

A interação social é uma parte essencial e onipresente da vida humana, além disso, a cognição social (mecanismos que nos permitem compreender os outros) inclui o conhecimento necessário para a interação e a formação de relações sociais. O quanto essas interações, bem como o processo de seletividade em compreender e/ou filtrar essas informações podem relacionar-se entre si, e na obtenção da chamada interação social, o quanto isso pode estar comodulado, podem resultar em estímulos atendidos ou em grande parte, o quanto estes estímulos podem estar inconscientes ou não.

Para explorar a frequência cardíaca e comportamental da interação social, este projeto de pesquisa tem como objetivo avaliar a interação entre dois ou mais indivíduos. Neste trabalho, revisamos estudos sobre mecanismos neurais e marcadores de interações sociais, entretanto o viés da pesquisa foi embasado em realizar um experimento comportamental via análise da variabilidade da frequência cardíaca.

Assim, tendo como base este arcabouço entre interfaces citadas, isto é, entre a frequência cardíaca e a linguagem, este trabalho tem como objetivo propor um estudo das interações sociais via análise da variabilidade da frequência cardíaca, empregando técnicas semelhantes às descritas por Dikker et al. (2014). Outrossim, este trabalho apresenta outras quatro seções, além desta introdução. A saber, a segunda seção traz uma breve fundamentação teórica, baseada na revisão da literatura, com a conceituação de interações sociais e predição. A terceira, descreve-se a metodologia do trabalho. Na sequência, na quarta seção, são apresentadas as atividades desenvolvidas, e na quinta seção, os resultados que se esperam. A sexta seção apresenta as referências bibliográficas.

1. **Fundamentação teórica e literatura empírica**

Pesquisas recentes mostraram o grau em que falantes e ouvintes, exibem padrões semelhantes de atividade cerebral durante a interação linguística humana, sendo esta, correlacionada com o sucesso comunicativo obtido através da natureza altamente preditiva.

Cérebros, em sua estrutura, possuem feixes de células que suportam a percepção e a ação, tentando constantemente combinar as entradas sensoriais com as expectativas ou previsões de cima para baixo, em que visa minimizar o erro de previsão dentro de uma cascata bidirecional de processamento cortical (Clark, A. 2013). Ao antecipar eventos, ao nosso redor, pode-se preparar respostas comportamentais, sendo rápidas e direcionadas, melhorando o isolamento e a identificação de sinais relevantes em dado ambiente e engajando-se em uma compreensão rápida e eficiente da linguagem (Maess et al., 2016).

Há um consenso de que as previsões baseadas em contexto facilitam o processamento léxico-semântico, com esses dados, modelos internos são gerados em antecipação aos atos de fala tanto na compreensão quanto na produção da linguagem, levando a uma atividade cerebral relativamente maior à medida que a previsibilidade aumenta: a preparação de atos da fala altamente previsíveis foi proposta para aumentar o ganho de atenção para suas consequências perceptivas esperadas em produção de linguagem, e percepções/palavras previsíveis são indiscutivelmente pré-ativadas, antes de serem vistas ou ouvidas durante a compreensão da linguagem (Dikker et al., 2014). Nesse contexto, as previsões no processamento da linguagem foram relatadas para muitos níveis linguísticos, como semântica lexical, sintática e sentencial, onde além disso, informações contextuais têm um impacto imediato nas previsões linguísticas, de acordo com (Hale, 2001).

Nesse consenso, fatores sintáticos podem afetar as respostas comportamentais e neurais durante o processamento da linguagem, no entanto, os mecanismos que permitem essa rápida extração de informações sintaticamente relevantes permanecem pouco compreendida. Durante o processo da compreensão falada ou escrita, cada palavra é totalmente analisada e interpretada em seu contexto. A compreensão da linguagem é um processo exigente, que requer a decodificação de sinais de fala altamente estruturados em um tempo muito curto. O sucesso da percepção da fala é aumentado através da combinação de análise orientada por entrada e baseada em previsão.

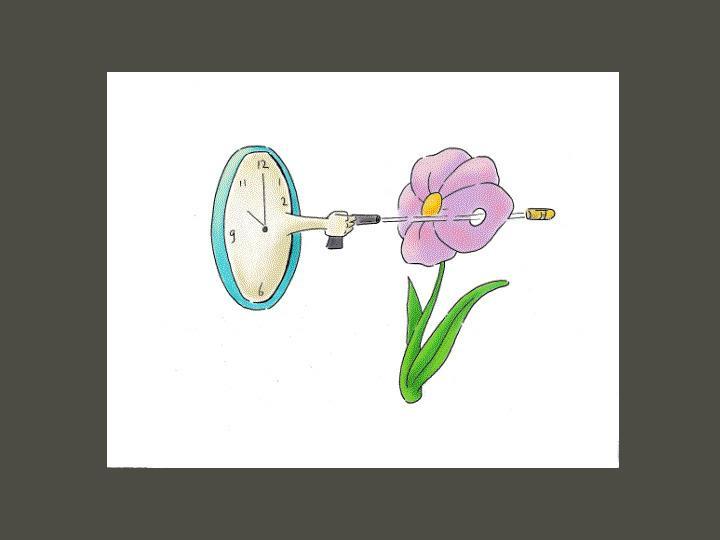
Entretanto, já as previsões são derivadas tanto do conhecimento prévio quanto da informação contextual. De acordo com a teoria da codificação preditiva, nosso cérebro antecipa continuamente a entrada sensorial atual, transferindo informações de áreas hierarquicamente mais altas para áreas mais baixas por meio de processamento de cima para baixo. Isso reduz as demandas de processamento em níveis mais baixos de hierarquia, isso se a entrada sensorial corresponder às expectativas. Do contrário, se a previsão estiver errada, o desempenho será inferior ao ideal, pois as ações apropriadas para a entrada prevista podem já ter iniciado, conforme (Friston, 2003).

1. **Metodologia**

Neste trabalho, foi realizado o experimento comportamental descrito por Dikker [1], no qual foi feito um experimento piloto com 5 aquisições, um orador e um ouvinte para validar o pipeline de análise. No entanto, para participar, os voluntários tinham que apresentar visão normal ou corrigida em relação ao normal, e sem nenhum histórico de distúrbios psiquiátricos ou neurológicos, livres de qualquer medicação psicotrópica.

Durante o experimento, foram apresentadas 45 imagens coloridas desenhadas à mão retratando cenas fictícias em que um animal ou objeto realiza uma ação sobre outro animal ou objeto (por exemplo, um pinguim abraçando uma estrela), tais cenas foram construídas com base em frases que serão criadas através da combinação aleatória de 45 verbos transitivos e 90 substantivos, denotando objetos, animais e alimentos comuns.

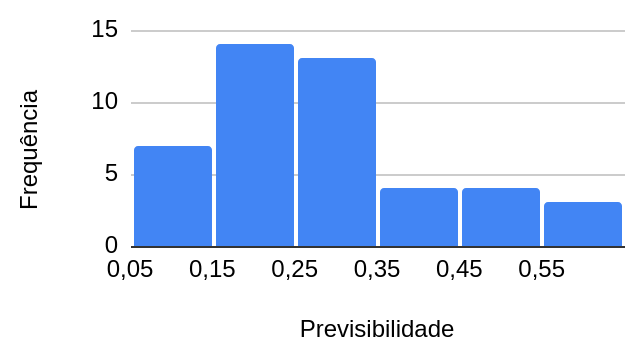
**Figura 1** - Imagem utilizada no experimento



Fonte: Dikker, Suzanne, et al. (2014).

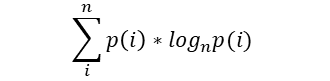
No experimento *online*, a cada imagem foi atribuída uma pontuação de previsibilidade, derivada de um questionário *online* no qual 42 voluntários descreveram, por escrito, cada uma das 45 cenas com a descrição que consideraram mais apropriada, além disso, para cada cena, atribuíram uma pontuação, em uma escala de 1 a 5 o quão certos estavam de que outras pessoas também inseririam exatamente a mesma frase, ademais nenhum desses voluntários participaram do experimento presencial. A previsibilidade é dada por 1 menos o valor da entropia, sendo que 0.0 é menos previsível e 1.0 mais previsível. Com base na distribuição da previsibilidade pelos itens, os itens serão atribuídos a uma de duas condições: alta previsibilidade e baixa previsibilidade.

A Figura 2 demonstra o histograma construído a partir do coeficiente de previsibilidade calculado com base no experimento online pelo método da entropia.

**Figura 2** - Frequência da previsibilidade. 

Fonte: elaborado pelos autores.

O método da entropia baseia-se na desordem do sistema, assim como o objetivo foi calcular a previsibilidade das frases, adotamos o cálculo complementar desse valor, no qual foi subtraído de 1 o valor da entropia, podendo variar de zero a 1, definido por:

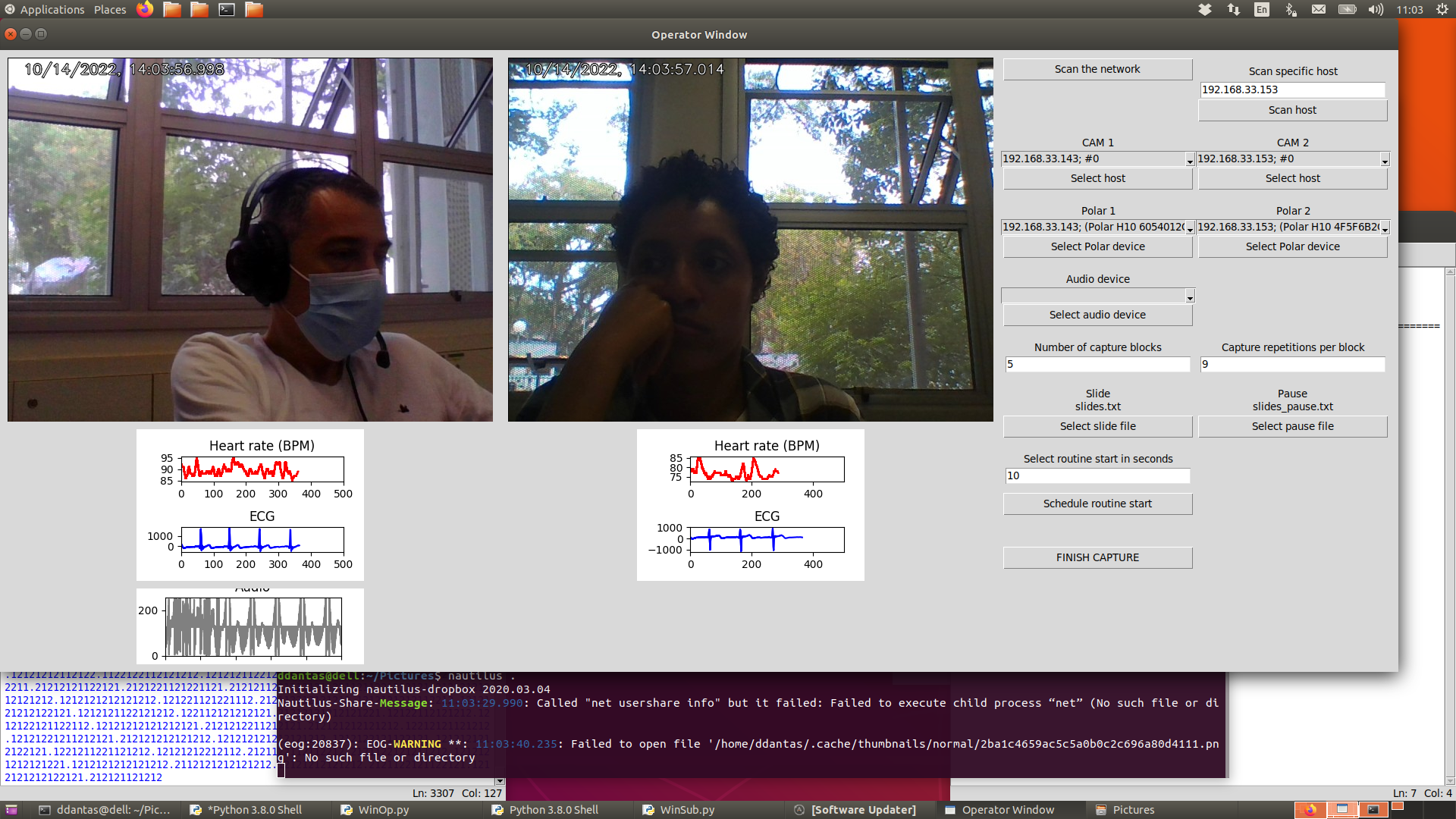


No experimento presencial, as sentenças serão descritas pelos participantes oradores, no qual foi instruído a descrever as imagens, usando frases declarativas, simples no tempo presente contínuo, com um único verbo transitivo e sem adjetivos ou frases adverbiais, durante as sessões do experimento presencial, e simultaneamente serão escutadas pelos ouvintes durante a captura. Em seguida, tanto para o orador quanto para os ouvintes, apresentaremos cada imagem durante 7,5 segundos, seguida de 7,5 segundos em tela cinza e depois de cruzes de fixação intermitentes (375 ms on/off, 3 segundos no total).

Em seguida dá-se início a exibição da próxima imagem. Cada participante verá um total de 45 ensaios em ordem aleatória, distribuídos em cinco blocos. Cada sessão durará aproximadamente 36 minutos e em paralelo será feita a captura de vídeo, áudio e frequência cardíaca por sensores de eletrocardiograma Polar H10.

1. **Resultados**

A realização da pesquisa exigiu a construção de um software capaz de processar capturas de frequências cardíacas dos participantes, nos experimentos presenciais, simultaneamente em paralelo a apresentação das imagens e captura das descrições das mesmas pelo voluntário orador. O programa conta com uma janela de controle do experimento, intitulada janela do operador, no qual é possível dar início da captura dos dados permitindo estabelecer tempo de exibição das imagens, além disso, a aplicação está programada para localizar automaticamente servidores de câmaras e de microfones na rede com uma determinada porta aberta, sensores (Polar H10) e eletrocardiograma disponíveis no localhost e nos demais hosts da rede, tal qual pausar e finalizar a captura. O sinal de eletrocardiograma, vídeo e ordem de exibição das imagens que ocorre com aleatoriedade são armazenados no computador de onde o experimento é manipulado.

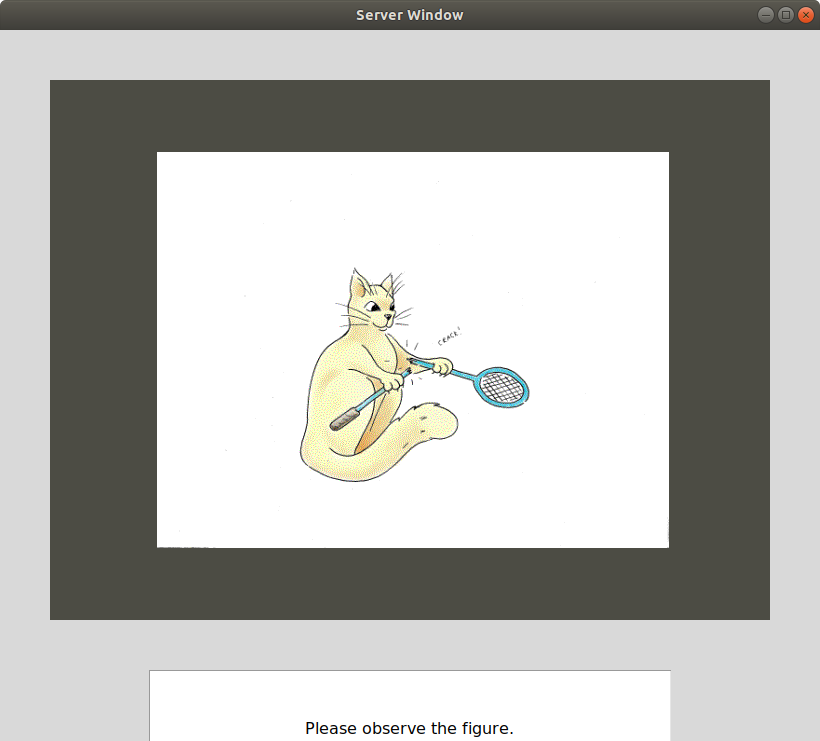
**Figura 3** - Janela do operador (autora e supervisor do projeto). 

Fonte: elaborado pelos autores.

A Figura 3 mostra a captura da janela do operador que controla a aquisição dos dados no experimento presencial, onde tem-se os displays de vídeos, de sinais (ECG e frequência cardíaca e barra de utilidades.

A coleta dispôs de três computadores, um deles executava um *Operator* enquanto os outros dois o *Subject*, estando todos conectados numa mesma rede. Cada sessão reunia dois voluntários (orador e ouvinte) que eram colocados numa sala, sentados lado a lado, com um computador em suas frentes. Para os participantes eram apresentadas somente imagens e mensagens de instrução.

**Figura 4** - Janela dos voluntários.



Fonte: elaborado pelos autores.

A Figura 4 exibe a janela que contém display de imagens e a caixa de mensagens direcionadas especificamente aos voluntários.

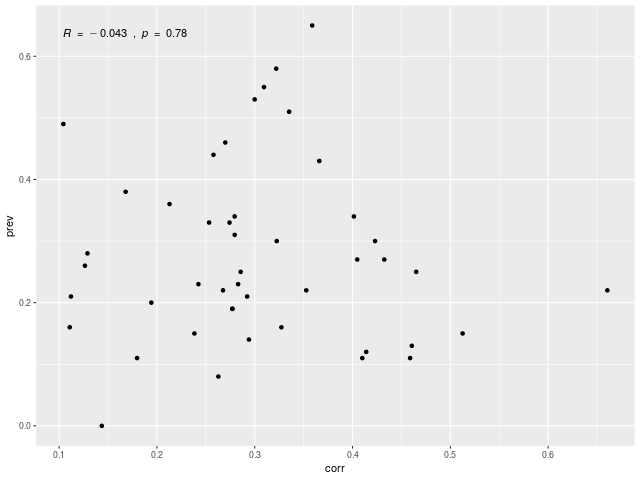
Ao final de uma captura, são gerados para cada participante um arquivo de vídeo contendo a gravação da câmera do participante (subj\*.mp4), um arquivo de dados ECG obtido pelos sensores Polar H10 (subj\* ecg.tsv) e um arquivo de dados de intervalo RR, também obtido por meio do sensor (subj\* rr.tsv). Após o tratamento dos dados, o resultado é um único arquivo .tsv com as seguintes colunas:

**Tabela 1** - Colunas do arquivo .tsv gerado.

| **CAMPOS** | **TIPO VALOR** |
| --- | --- |
| folder | String |
| Time | Integer |
| Block | Pause, Slide |
| Slide | Pause, 1, 2, …, 47 |
| hr\_subj\*\_linear | Float |
| hr\_subj\*\_nearest | Float |
| hr\_subj\*\_ecg\_linear | Float |
| hr\_subj\*\_ecg\_nearest | Float |
| d\_hr\_subj\*\_linear | Float |
| d\_hr\_subj\*\_nearest | Float |
| d\_hr\_subj\*\_ecg\_linear | Float |
| d\_hr\_subj\*\_ecg\_nearest | Float |
| rr\_subj\_linear | float |
| rr\_subj\_nearest | float |
| rr\_subj\_ecg\_linear | float |
| rr\_subj\_ecg\_nearest | float |

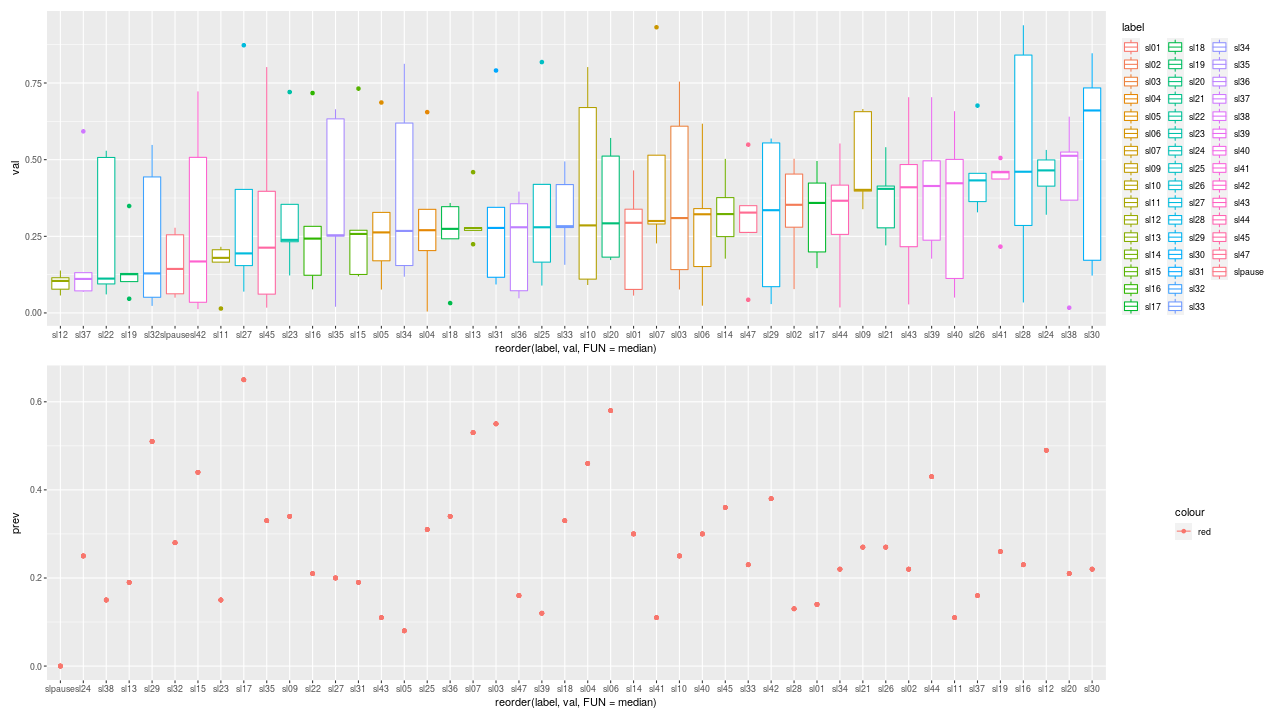
Fonte: elaborado pelos autores.

Conforme exposto acima, realizamos uma análise de correlação entre sujeitos para identificar a variabilidade da frequência cardíaca onde a atividade correlacionada entre o falante e os ouvintes fosse afetada pela previsibilidade.

**Figura 5** - Correlação entre frequência cardíaca e previsibilidade das frases, eixo x representando a correlação média dos batimentos e eixo y previsibilidade do slide

Então, para explicar os padrões de correlação observados, examinamos os dados, através do gráfico de dispersão relacionando os batimentos cardíacos e a previsibilidade das frases dos slides. Como resultado, temos que a correlação é essencialmente zero, desse modo não é possível afirmar que há relação entre variabilidade do ritmo cardíaco com a previsibilidade das frases, que seria o meio de interação social proposto na pesquisa. Ao plotar os resultados em *box-plot* de modo a avaliar a variabilidade da frequência cardíaca das capturas realizadas com o intuito de isolar e apresentar os slides comparando com o gráfico de previsibilidade plotado a partir dos resultados do experimento online, temos como resultado que não encontramos relação entre os parâmetros analisados, indicando que para esse experimento proposto em que é estruturado um ambiente de interação interpessoal entre duas pessoas controlado, não foi possível observar correlação.

**Figura 6** - Gráfico de *box-plots* com a distribuição das correlações entre os batimentos do subj 1 orador e subj2 ouvinte, ordenados pela correlação mediana e relaciona com o gráfico de pontos que plota previsibilidade dos slides



A figura 6 mostra se existe alguma tendência de crescimento ou decréscimo da previsibilidade em função do crescimento da correlação Porém, como é possível analisar não existe tendência aparente. Embasando a afirmação de que dado o ambiente controlado proposto no experimento não foi possível correlacionar os parâmetros em questão.

1. **Referências Bibliográficas**
2. Dale, Anders M., Bruce Fischl, and Martin I. Sereno. "Cortical surface-based analysis: I. Segmentation and surface reconstruction." Neuroimage 9.2 (1999): 179-194.
3. Dikker, Suzanne, et al. "Early occipital sensitivity to syntactic category is based on form typicality." Psychological Science 21.5 (2010): 629-634.
4. Sänger, Johanna, Ulman Lindenberger, and Viktor Müller. "Interactive brains, social minds." Communicative & integrative biology 4.6 (2011): 655-663.
5. Bergerbest, Dafna, Dara G. Ghahremani, and John DE Gabrieli. "Neural correlates of auditory repetition priming: reduced fMRI activation in the auditory cortex." Journal of Cognitive Neuroscience 16.6 (2004): 966-977.
6. Desimone, Robert, and John Duncan. "Neural mechanisms of selective visual attention." Annual review of neuroscience 18.1 (1995): 193-222.
7. Clark, Andy. "Whatever next? Predictive brains, situated agents, and the future of cognitive science." Behavioral and brain sciences 36.3 (2013): 181-204.
8. Maess, Burkhard, et al. "Prediction signatures in the brain: semantic pre-activation during language comprehension." Frontiers in Human Neuroscience 10 (2016): 591.
9. Dikker, Suzanne, et al. "On the same wavelength: predictable language enhances speaker–listener brain-to-brain synchrony in posterior superior temporal gyrus." Journal of Neuroscience 34.18 (2014): 6267-6272.
10. Friston, Karl. "Learning and inference in the brain." Neural Networks 16.9 (2003): 1325-1352.
11. Hale, John. "A probabilistic Earley parser as a psycholinguistic model." Second meeting of the north american chapter of the association for computational linguistics. 2001.

**Apêndice**

**Dados e colunas do arquivo resultante do pré-processamento**

Os dados coletados foram registrados pelo software em momentos de tempo de não inteiros (milisegundos). Todas as colunas que possuem valores de frequências cardíacas, foram realizadas interpolações para obter novos valores correspondentes a tempos inteiros. Algumas colunas tem dois campos, devido a captura ter sido feita em dois participantes, no qual pode assumir o valor de 1 ou 2.

**Descrição das colunas da tabela**

* folder: A cada iteração do experimento os arquivos gerados durante a captura são salvos em um diretório novo. Essa coluna indica a qual captura um determinado conjunto de dados pertence, explicitando o nome dos diretórios de onde os dados vieram;
* time: é o instante inteiro relacionado ao conjunto de dados de uma certa linha;
* block: indica qual é o bloco de captura atual (slide ou pause).
* slide: imagens (1 a 47) ou pause (tela cinza);
* hr\_subj\*\_linear: Valor de frequência cardíaca dos participantes obtido utilizando interpolação linear (1 e 2);
* hr\_subj\*\_nearest: Valor de frequência cardíaca dos participantes obtido utilizando a estratégia de “vizinho mais próximo" para determinar qual valor pertencia a qual segundo inteiro (1 e 2);
* hr\_subj\*\_ecg\_linear: Valores de frequência cardíaca inferidos a partir do sinal ECG coletado e interpolados do mesmo modo que hr\_subj\*\_linear (1 e 2);
* hr\_subj\*\_ecg\_nearest: Valores de frequência cardíaca inferidos a partir do sinal ECG coletado e interpolados do mesmo modo que hr\_subj\*\_nearest (1 e 2);
* d\_hr\_subj\*\_linear: Derivada dos valores da coluna hr\_subj\* linear (1 e 2);
* d\_hr\_subj\*\_nearest: Derivada dos valores da coluna hr\_subj\*\_nearest (1 e 2);
* d\_hr\_subj\*\_ecg\_linear: Derivada dos valores da coluna hr\_subj\*\_ecg\_linear (1 e 2);
* d\_hr\_subj\*\_ecg\_nearest: Derivada dos valores da coluna hr\_subj\*\_ecg\_nearest (1 e 2);
* rr\_subj\_linear: (1 e 2);
* rr\_subj\_nearest: (1 e 2);
* rr\_subj\_ecg\_linear: (1 e 2);
* rr\_subj\_ecg\_nearest: (1 e 2);