



Is the finger-voice span an  
indicator of reading  
proficiency?

**Vito Pirrelli**

ComPhys Lab  
Institute for Computational Linguistics  
National Research Council - Italy  
[vito.pirrelli@ilc.cnr.it](mailto:vito.pirrelli@ilc.cnr.it)

# gist of the talk

- finger-pointing reading requires the fine coordination of three motor systems with different intrinsic dynamics
  - voice **articulation**
  - **eye** movements
  - **finger** movements
- multimodal tracking evidence of their relative pace is a reliable indicator of a reader's decoding efficiency and language skills



finger-pointing reading

what is it?



- [finger-pointing reading](#) is based on the learners' ability to accurately pronounce a printed word in a text line while the index finger is pointing to it
- it is known to support the development of word reading in context in [young readers](#), through integration of the auditory perception of syllables and the visuo-spatial orientation of printed words built upon the visual and tactile exploration of their spatial dimension (Mesmer and Lake, 2010; Mesmer and Williams, 2015)

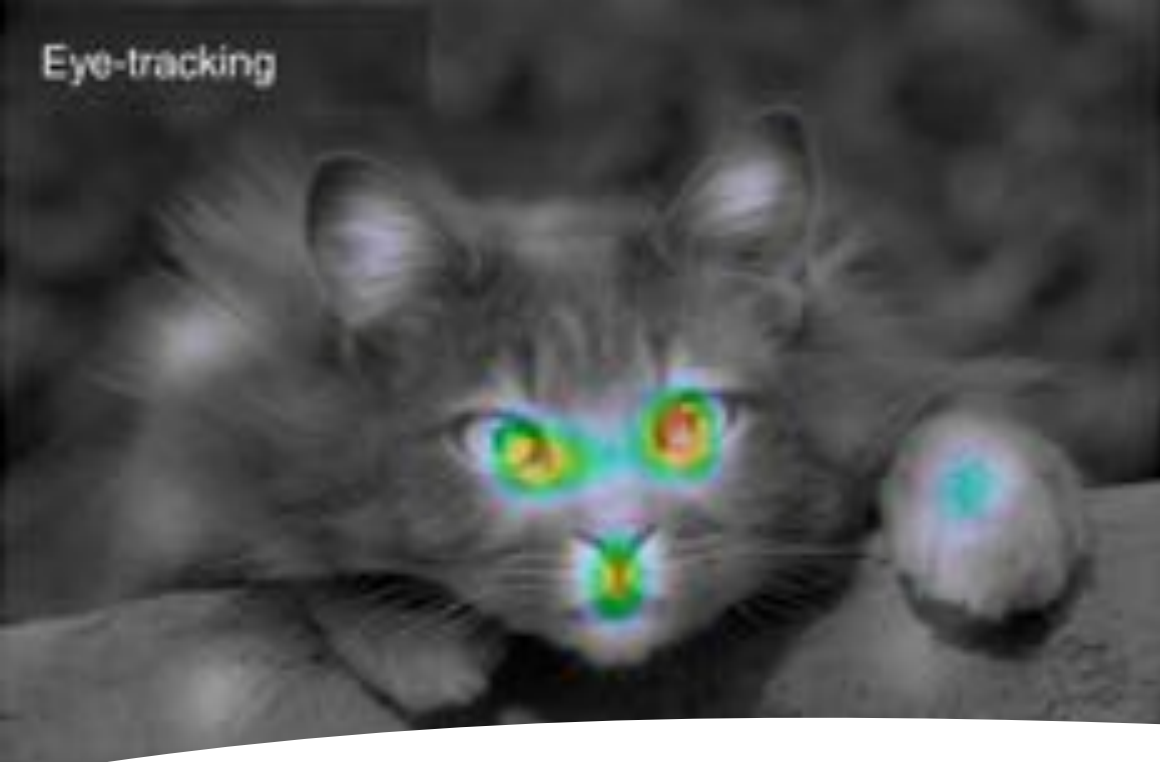


# the eye-finger connection (Lio et al. 2019)

---

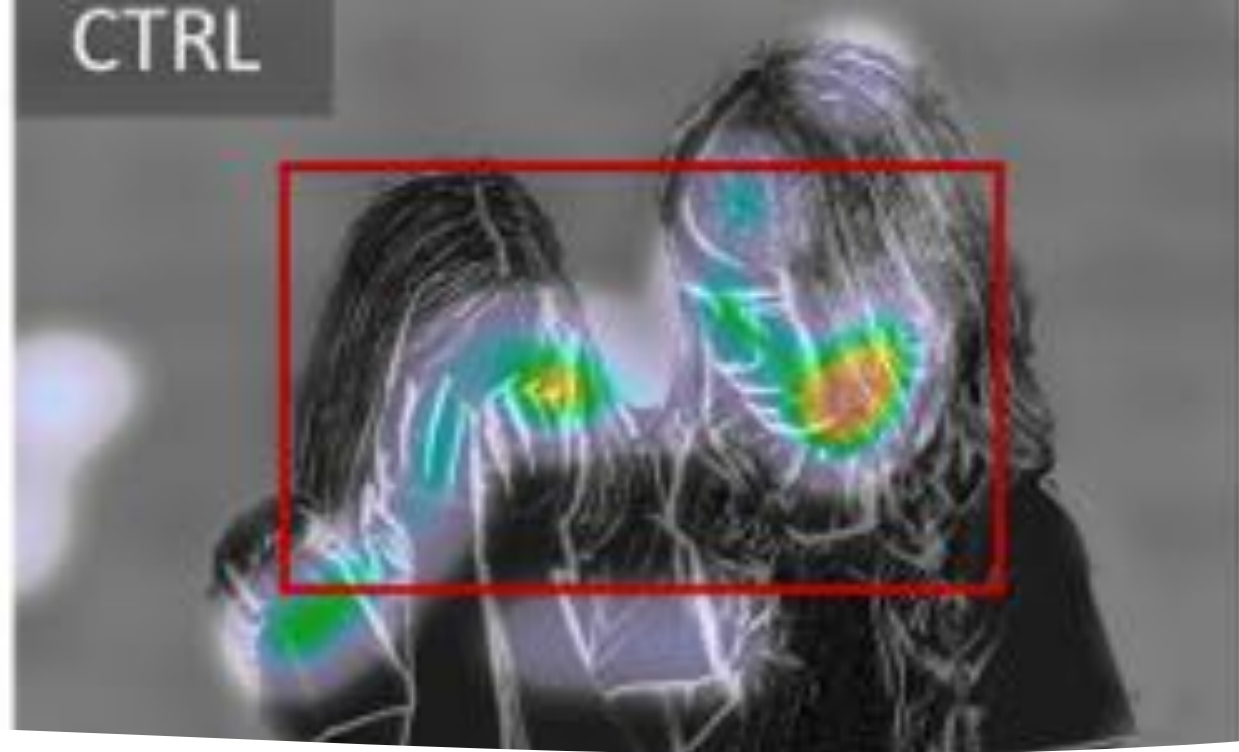
- deblurring





the eye-finger  
connection (II)

- evidence for a *strong correlation* between finger & eye movements in visual image exploration



the eye-finger  
connection (III)

- Autism Spectrum Disorder subjects vs controls

La prima obiezione che mi arriva è se raccontarle non possa generare emulazione . Che si tratti di resoconti giornalistici o di fiction , una parte di chi osserva finirà necessariamente per specchiarsi . E specchiarsi significa che non esiste emulazione possibile perché la strada intrapresa era già quella .



a text page as a  
visual space

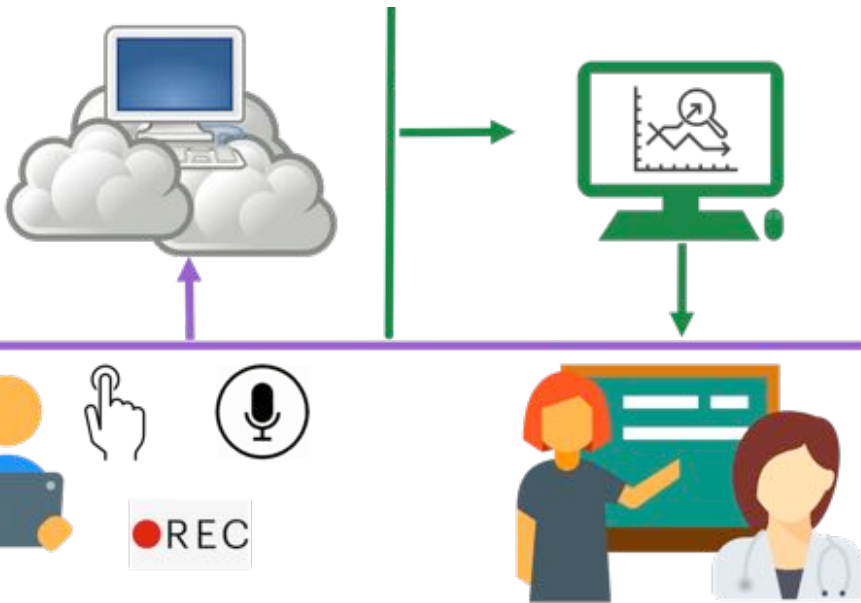




## ReadLet

finger-tracking everywhere

- a reader can use a **tablet** as a *reading book*
- while reading (silently or aloud), the reader is asked to **underline the words** she's reading with the finger of her dominant hand
- **finger movements and voice** are *captured* by the tablet, encrypted and sent to a central server through an *Internet connection*
- data are post-processed, **time-aligned** with the text and the reader's voice and analyzed in terms of *direction, path, speed*



# finger reading



finger-pointing reading

the finger as a  
proxy of the eye?



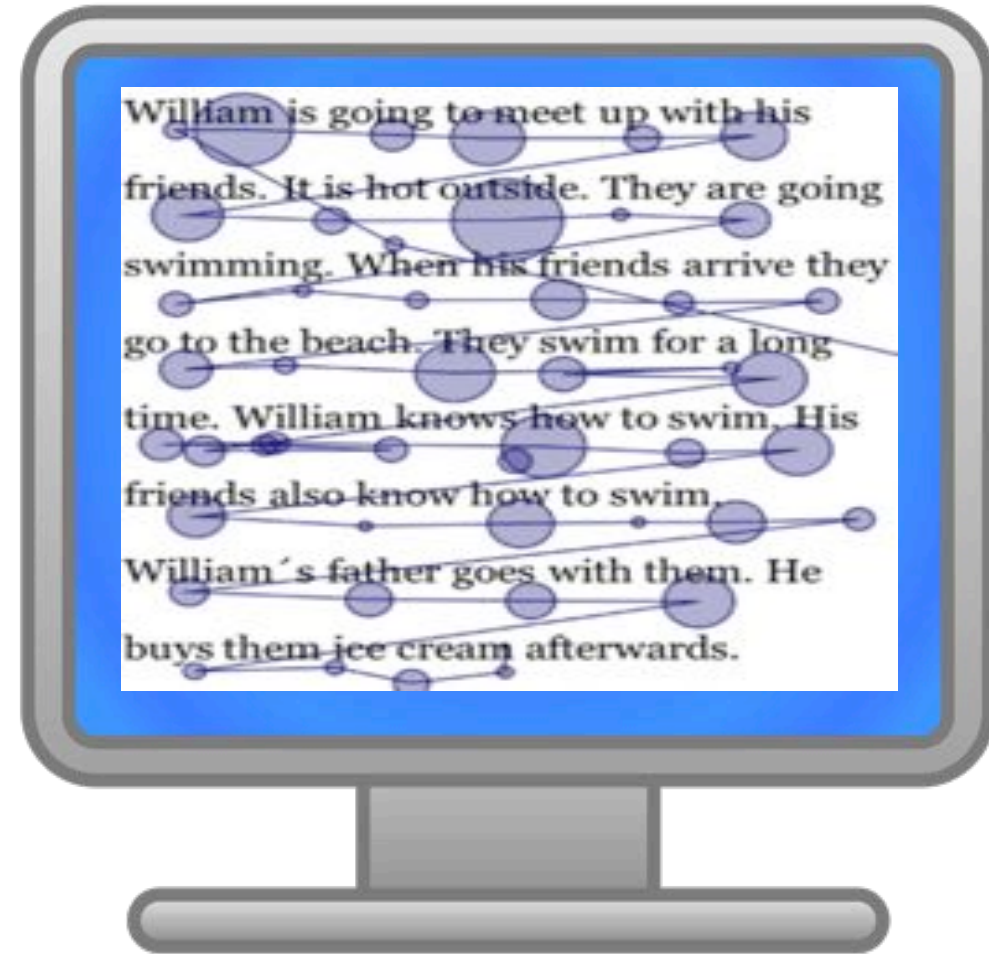


## eye tracking technology

- eye position
- eye movement
- millisecond precision

## eye reading

- discrete signal (saccade & fixations)
  - fixation duration
  - scanning path
  - fixation rate
  - skipping rate



# finger reading

- continuous signal (sliding movement)
  - tracking time
  - tracking path & direction
  - speed variation
  - (little skipping)



... Ziggy is an alien from a distant planet



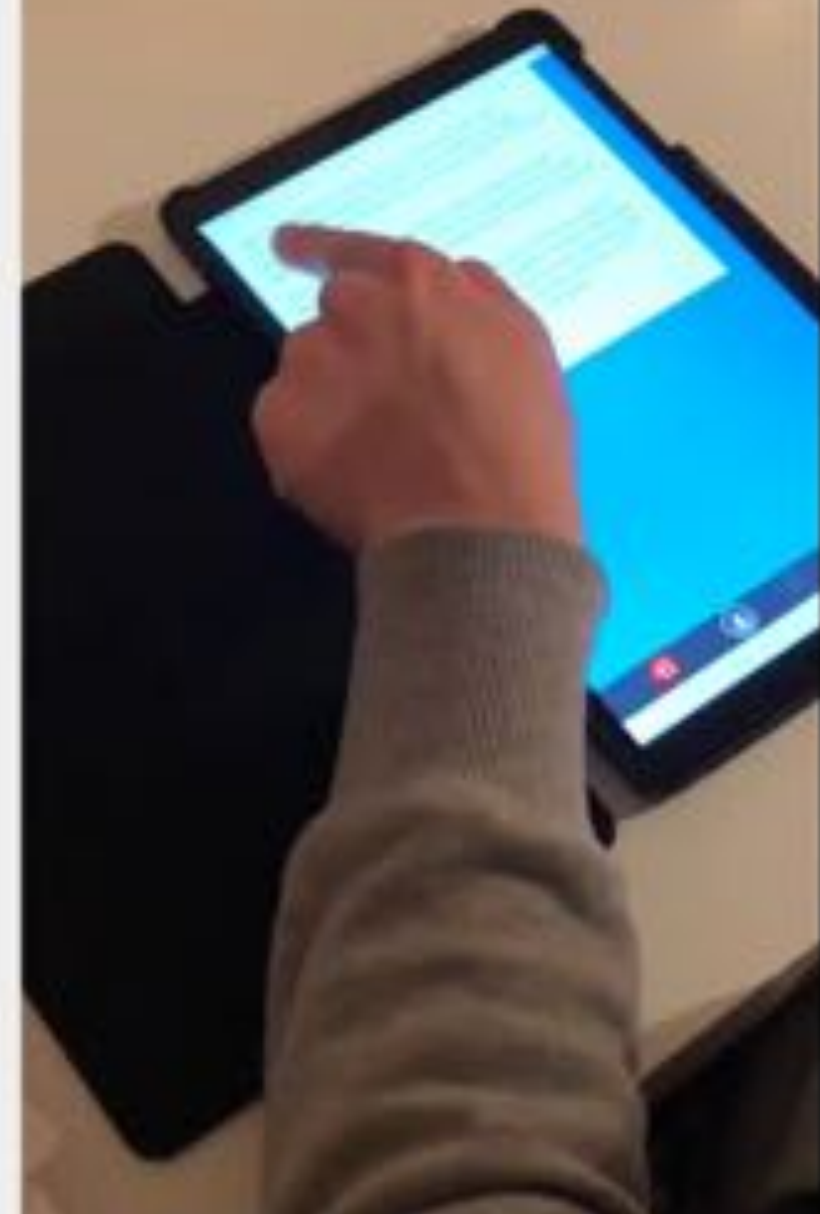
# finger reading

Linea (1.8x) 90-03-20 +6.200 (Nuovi) (px)

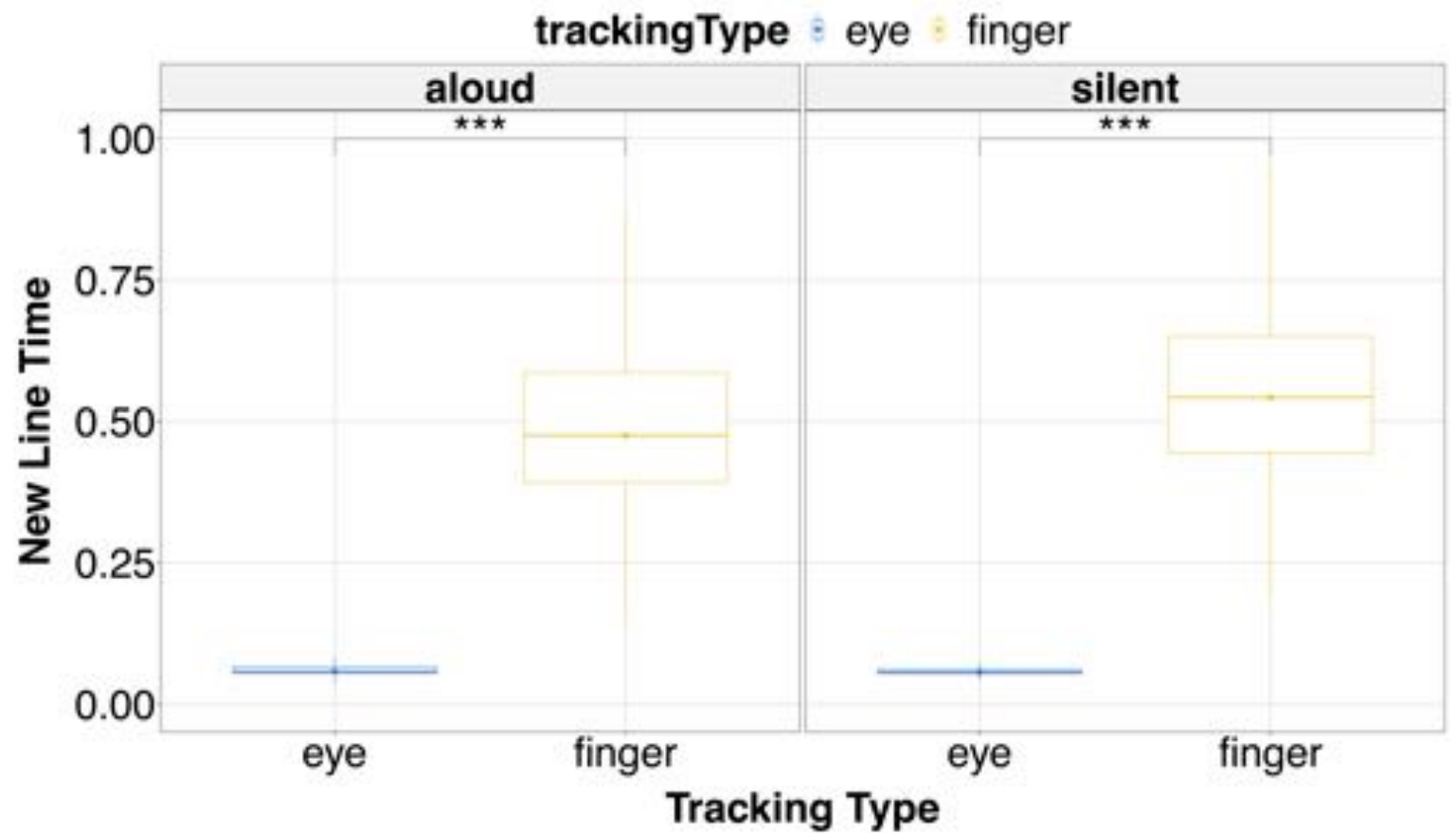
Nonno Andrea è molto bravo a raccontare le storie. Non racconta la storia della buonanotte. Il nonno Andrea non legge mai libri. Racconta storie che si vivono al momento. Mi piace molto ascoltare.

Con la sua voce dolce, nonno Andrea riesce a dare vita ad ogni parola. Mi sembra di vivere le storie che mi racconta. Le sue mani sono grandi e, mentre parla, si muovono veloci. Sembra un ballerino o un pittore in un teatro.

Sono molto felice di passare del tempo con lui. Prima di andare a dormire, lo mi sento per tenerlo accanto al letto e il nonno sulla sedia. Poi lui mi sorride e lo so che sta per cominciare. Per alcuni secondi chiude gli occhi e fa un grande respiro. E così cominciano tutti i viaggi, con la grande fantasia del nonno alla guida.

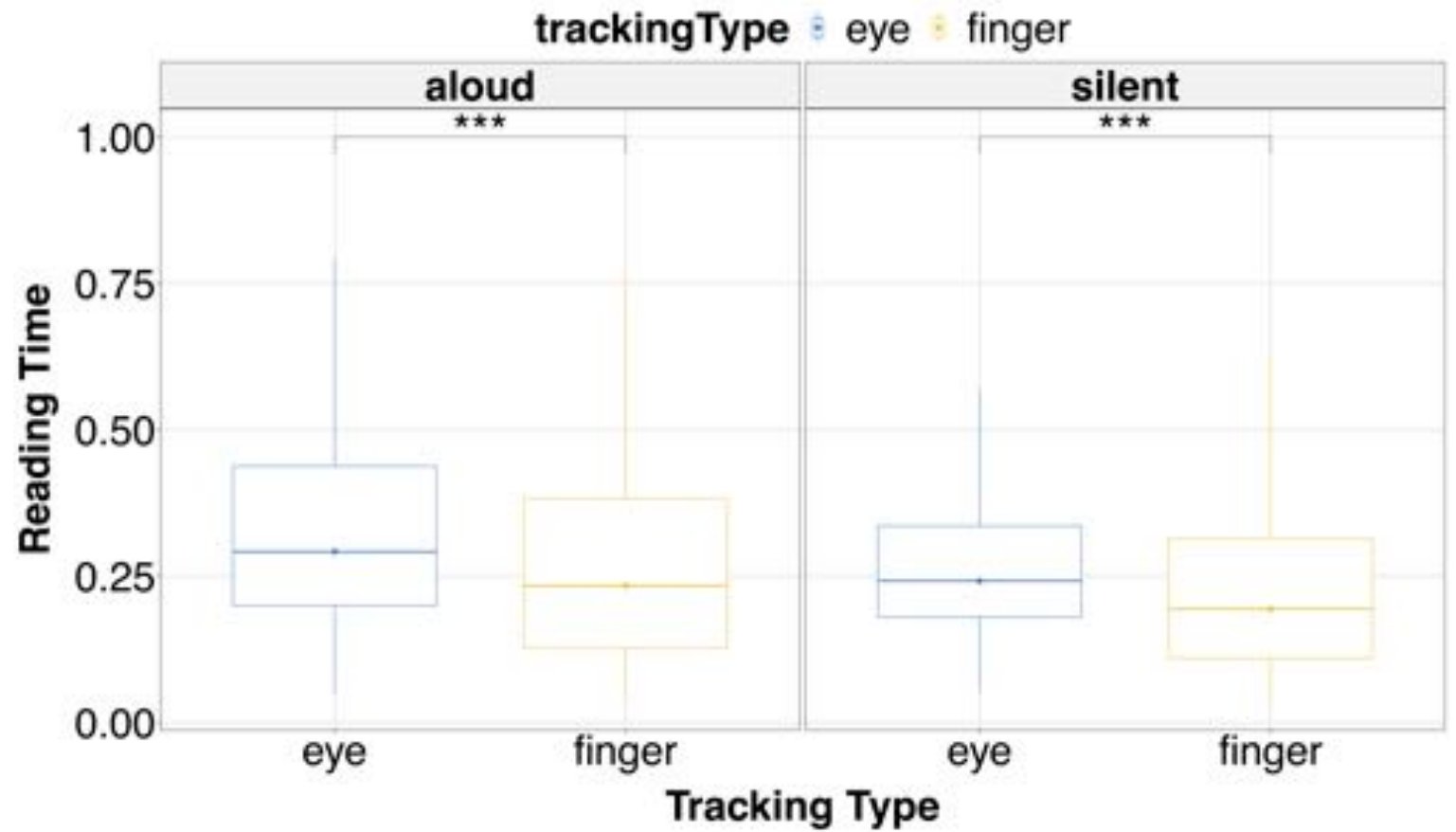


skipping  
to newline





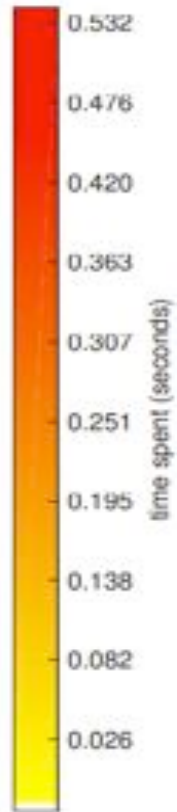
eye fixation >  
finger tracking



# The first objection that gets to me ...

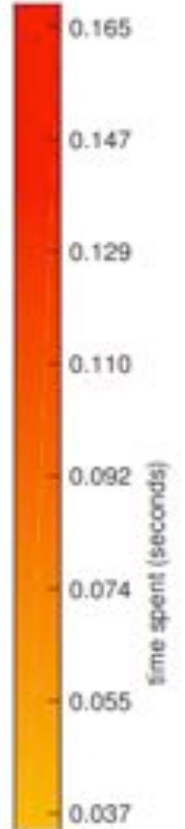
## eye tracking

La prima obiezione che mi arriva è se raccontarle non possa generare emulazione. Che si tratti di resoconti giornalistici o di fiction, una parte di chi osserva finirà necessariamente per specchiarsi. E specchiarsi significa che non esiste emulazione possibile perché la strada intrapresa era già quella.



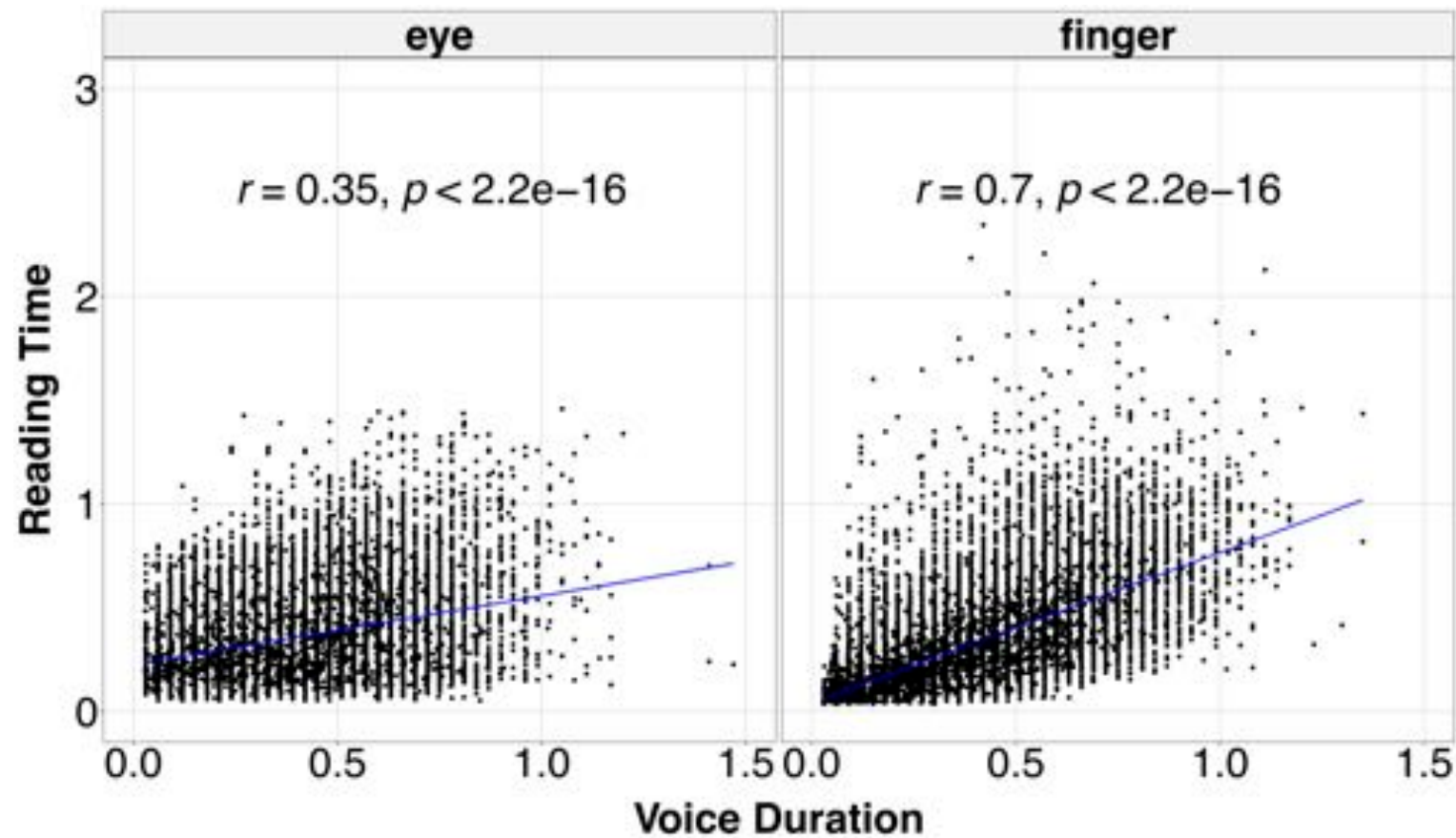
## finger tracking

La prima obiezione che mi arriva è se raccontarle non possa generare emulazione. Che si tratti di resoconti giornalistici o di fiction, una parte di chi osserva finirà necessariamente per specchiarsi. E specchiarsi significa che non esiste emulazione possibile perché la strada intrapresa era già quella.



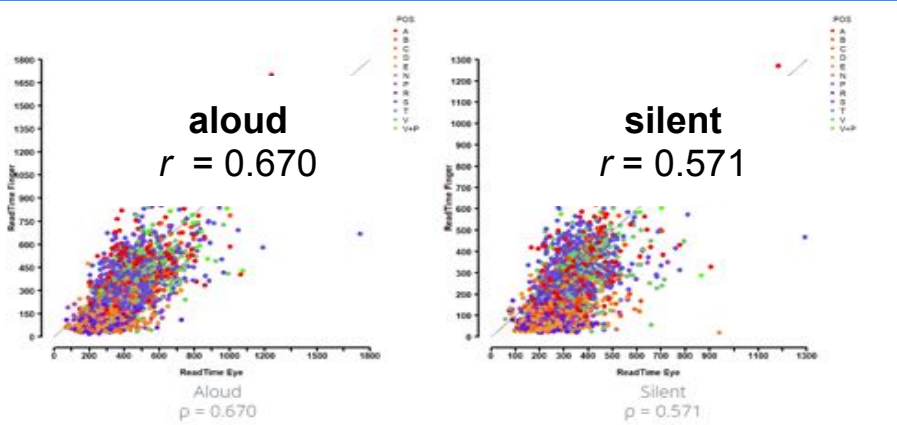
discrete vs. continuous

correlation  
with voice  
rate

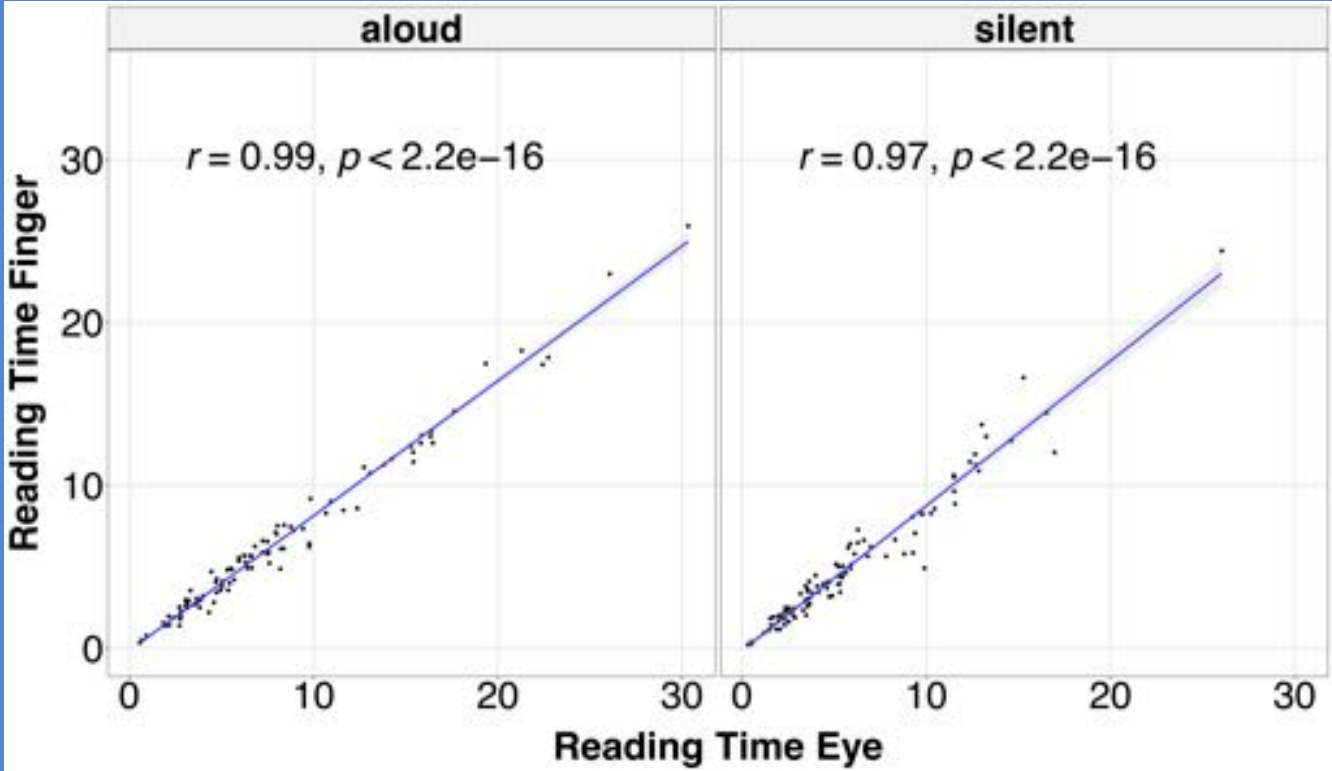


# eye-finger correlation

words



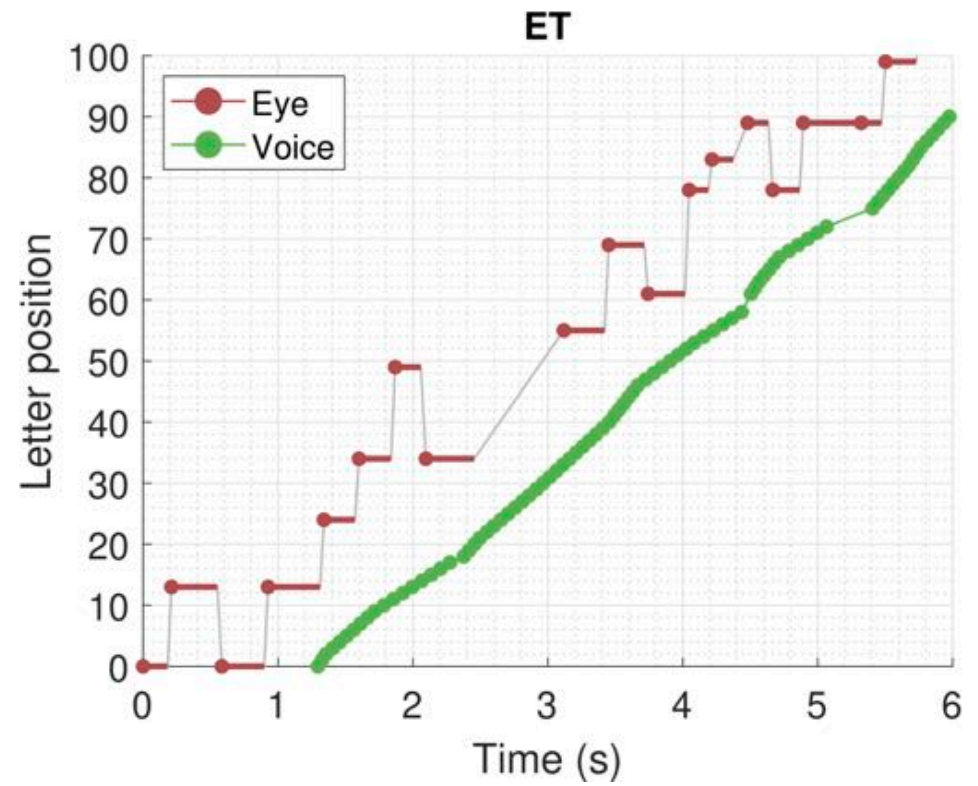
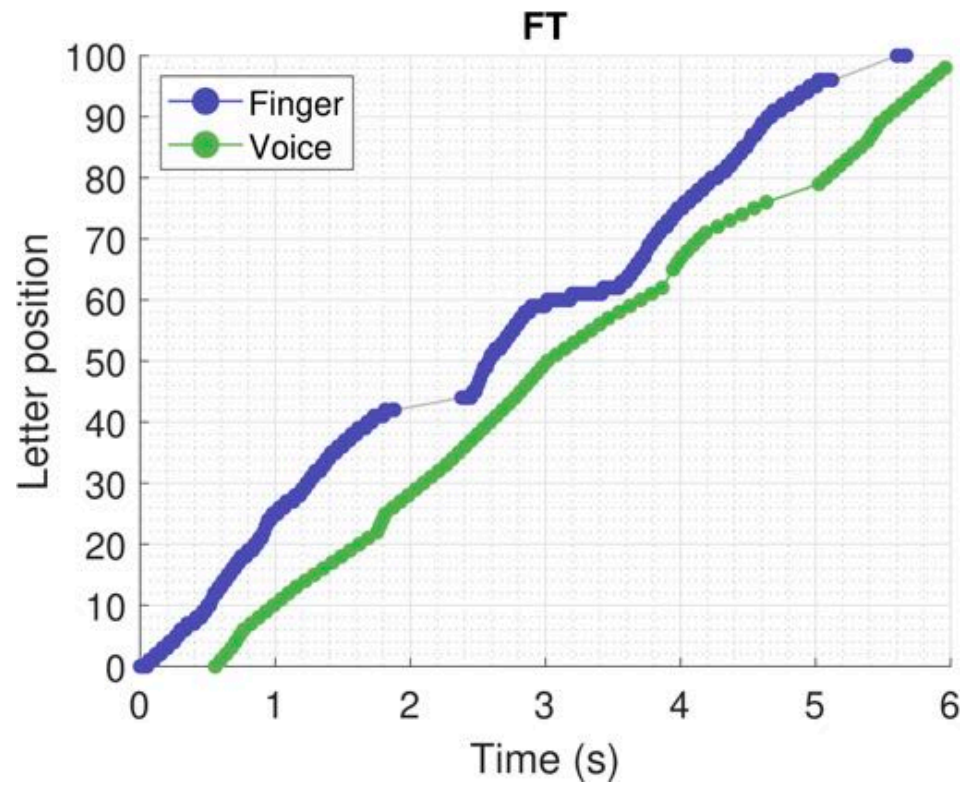
sentences



finger-pointing reading

eye-voice span  
finger-voice span





voice, finger and eye movements

time (1.0x): 00:00:05 +0.079 / 00:04:22 (NaN/25 fps)

"Dance On A Volcano" dei Genesis è un brano potentissimo che predispone l'animo a cose straordinarie. È la colonna sonora «del capolavoro del genio Makko», come dice Diego Bianchi, andato in onda a Propaganda Live il 5 aprile scorso.

Piccoli filamenti neri, potrebbero essere peli di ratto o anche fuliggine, convergono verso un punto a formare uno strano animale nero. Sono milioni di filamenti e presto l'animale diventa enorme. Davanti a lui un ragazzino, Simone, più piccolo della bocca rossa e spaventosa della bestia, sta lì, fermo. Saldo. Attorno a lui gli anziani si accalcano e gli dicono: «Vieni via ragazzo, quella bestia nera, quando si forma, è invincibile. L'unica è far finta di non vederla, è aspettare che passi. Fidati, noi la conosciamo bene».

DECODING (COMV\_V03)

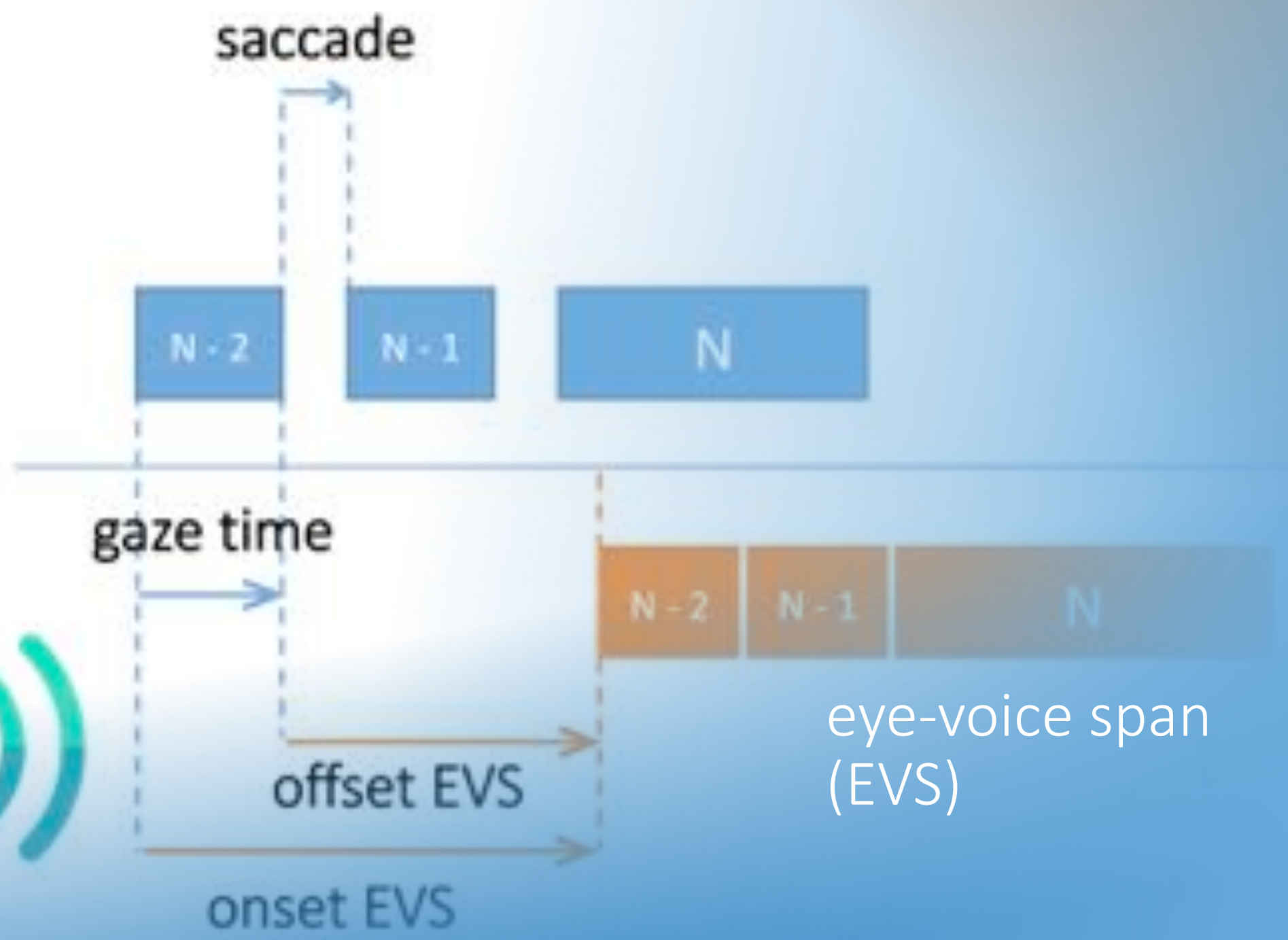


time (1.0x): 00:00:00 +0.000 / 00:04:37 (NaN/25 fps)

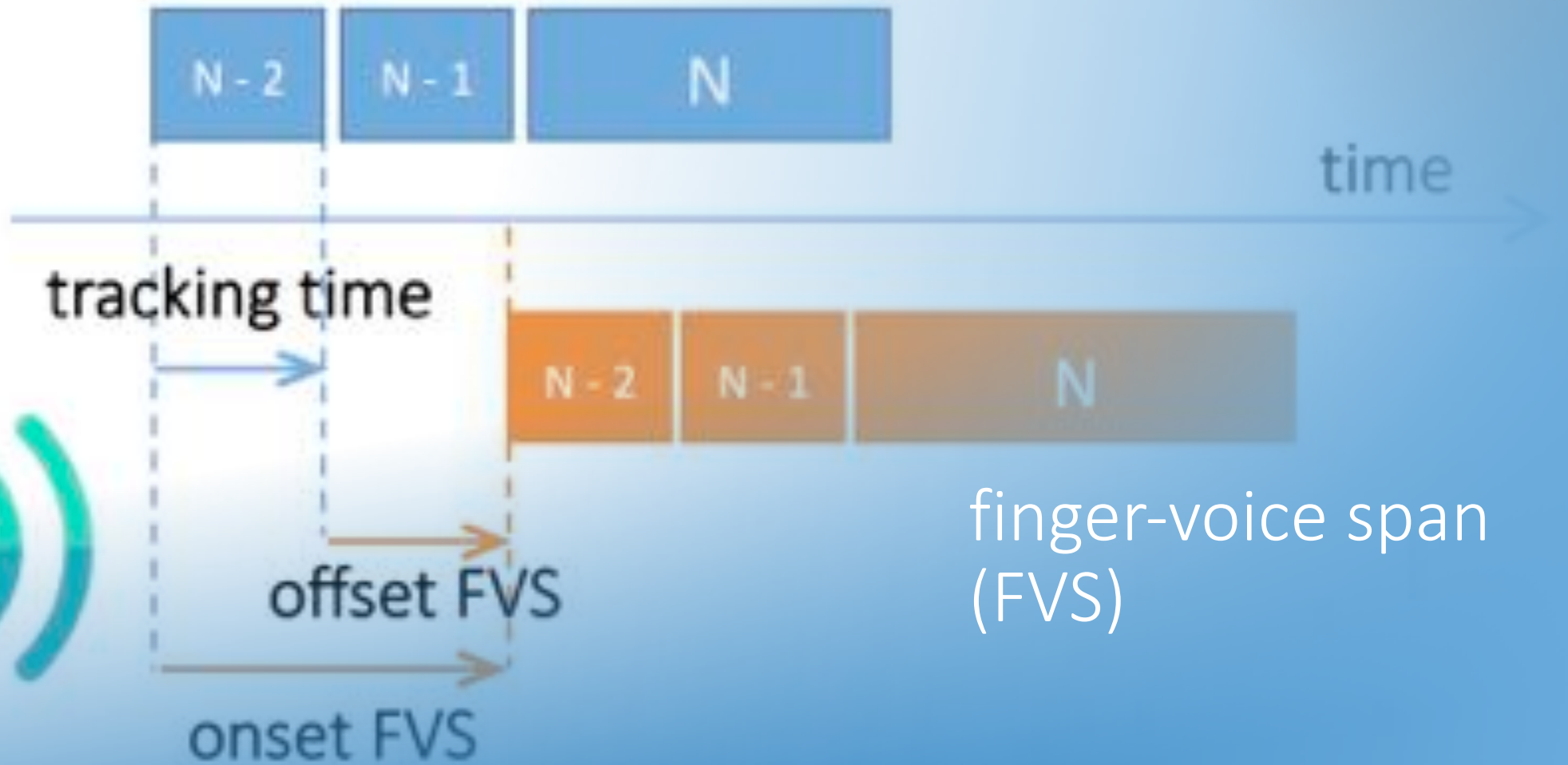
Per conoscere la storia di un paese non basta studiare quella delle sue istituzioni, ma è necessario anche sondare la parte sommersa, scavare tra gli atti giudiziari e cercare di venirne a capo. Le organizzazioni criminali sono consorzi criminal-imprenditoriali che non hanno sofferto la crisi economica, basando il loro core business sul monopolio del traffico di stupefacenti. Con i proventi di quello invadono ogni altro settore. Non raccontare questo meccanismo e non raccontare i protagonisti significa conoscere la storia a metà. Ecco perché ho deciso di portare in televisione le biografie dei boss.

La prima obiezione che mi arriva è se raccontarle non possa generare emulazione. Che si tratti di resoconti giornalistici o di fiction, una parte di chi osserva finirà necessariamente per specchiarsi. E specchiarsi significa che non esiste emulazione possibile perché la strada intrapresa era già quella.

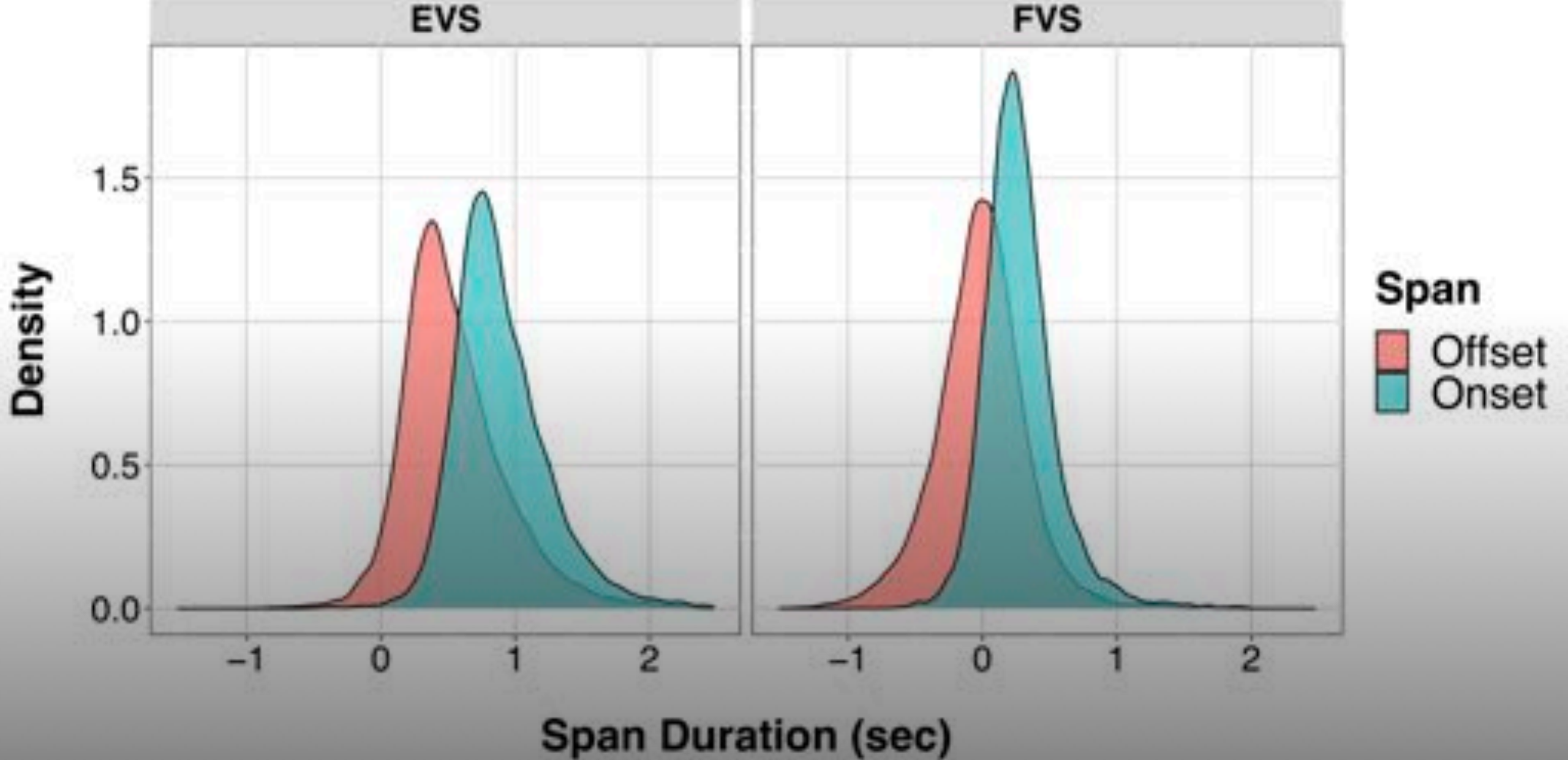




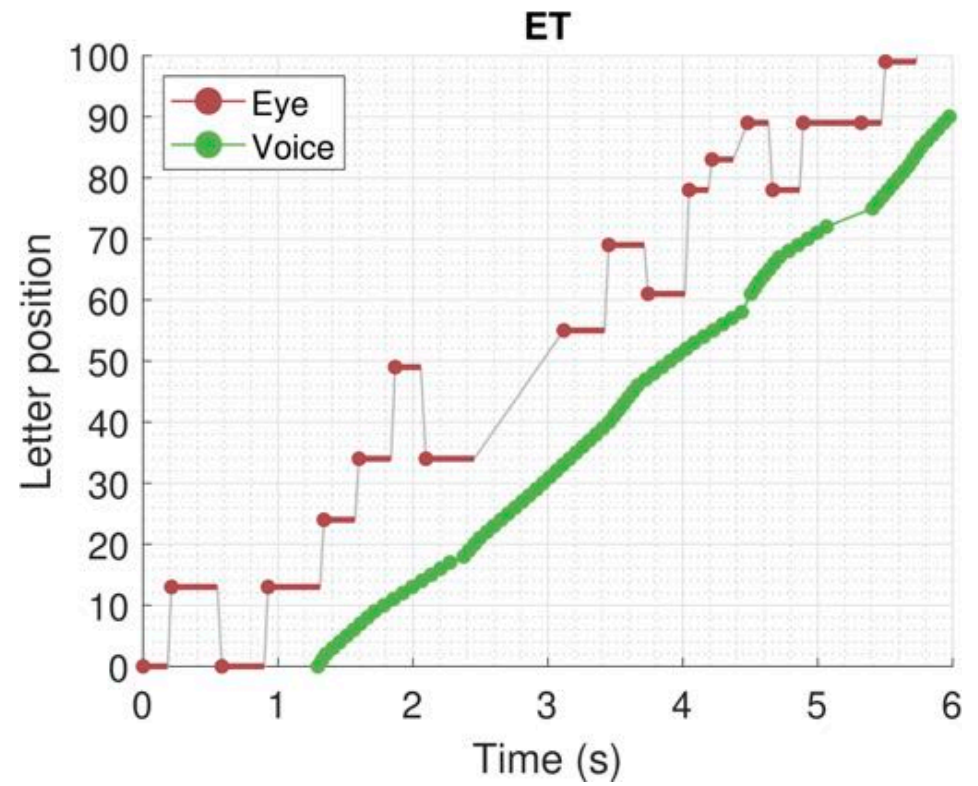
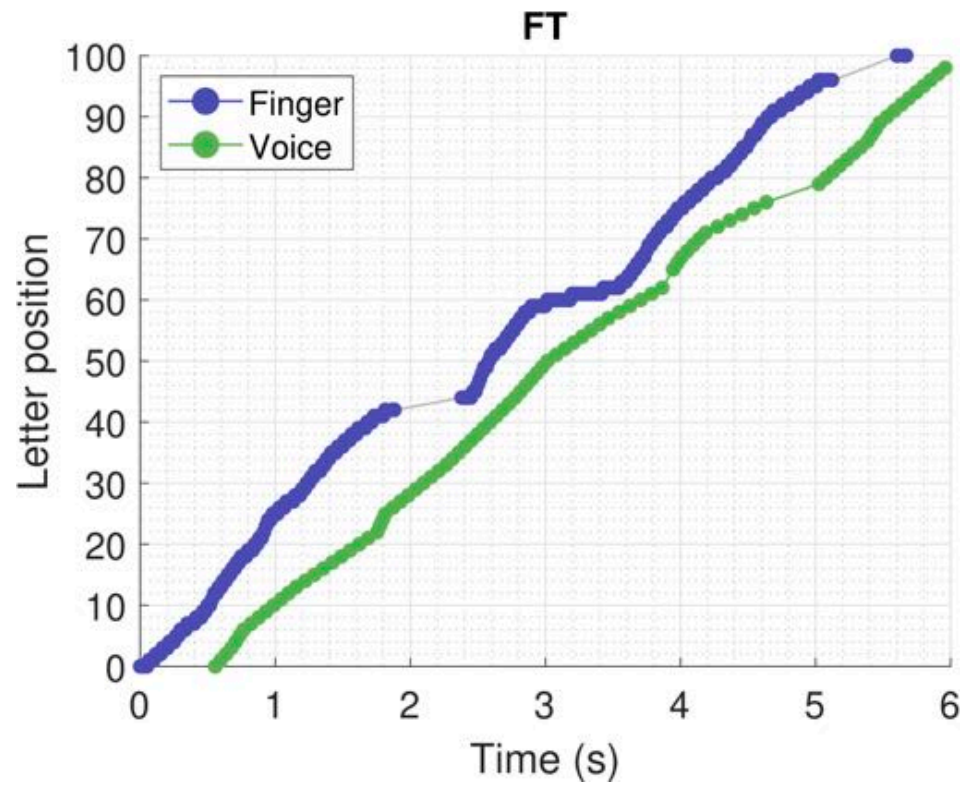
eye-voice span (EVS)



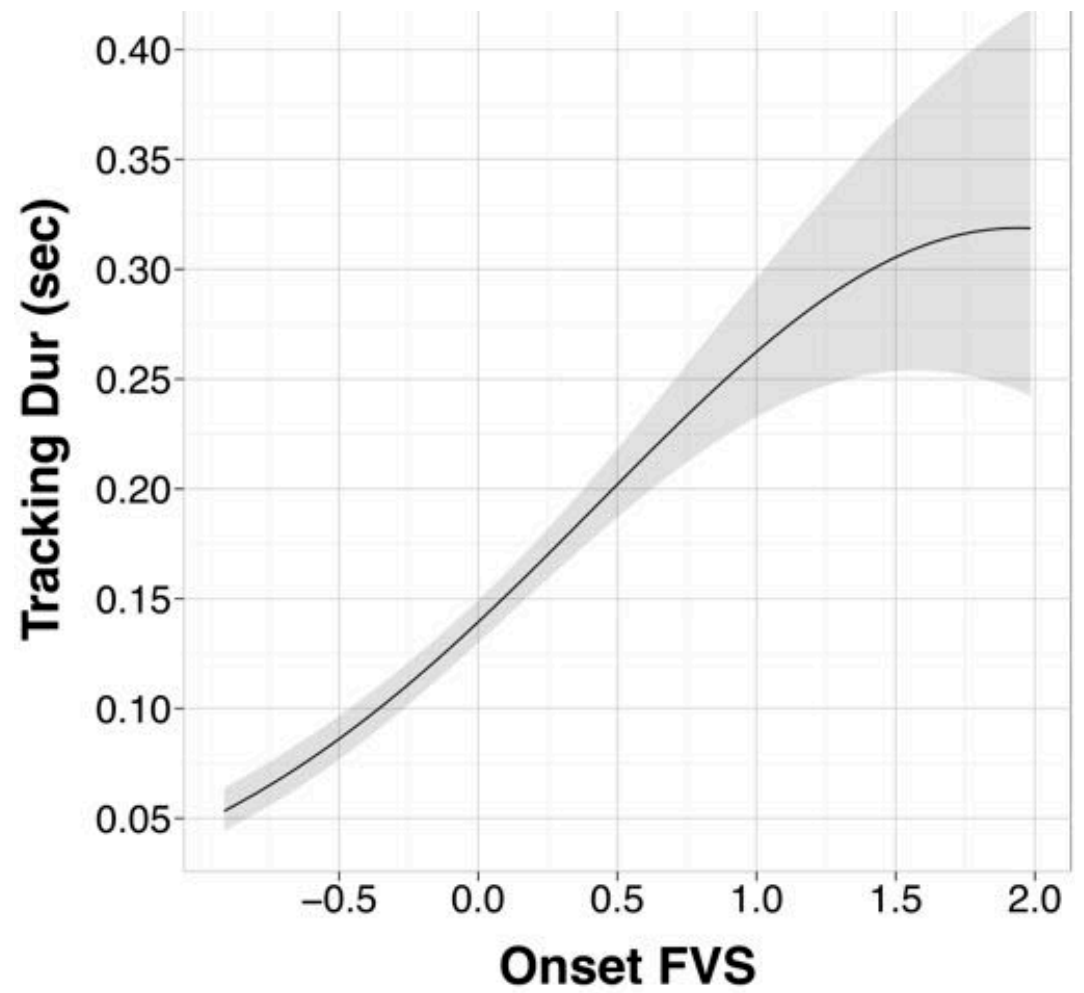
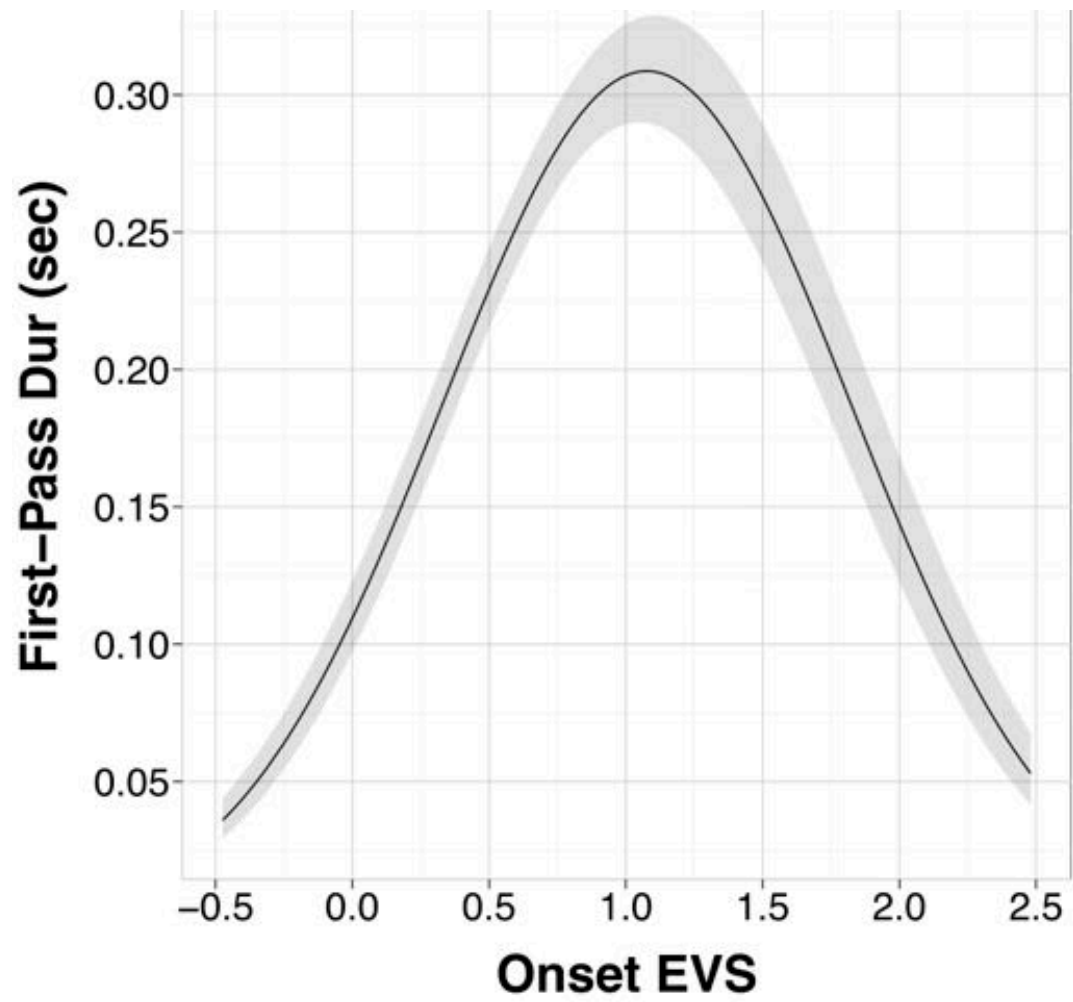
finger-voice span  
(FVS)



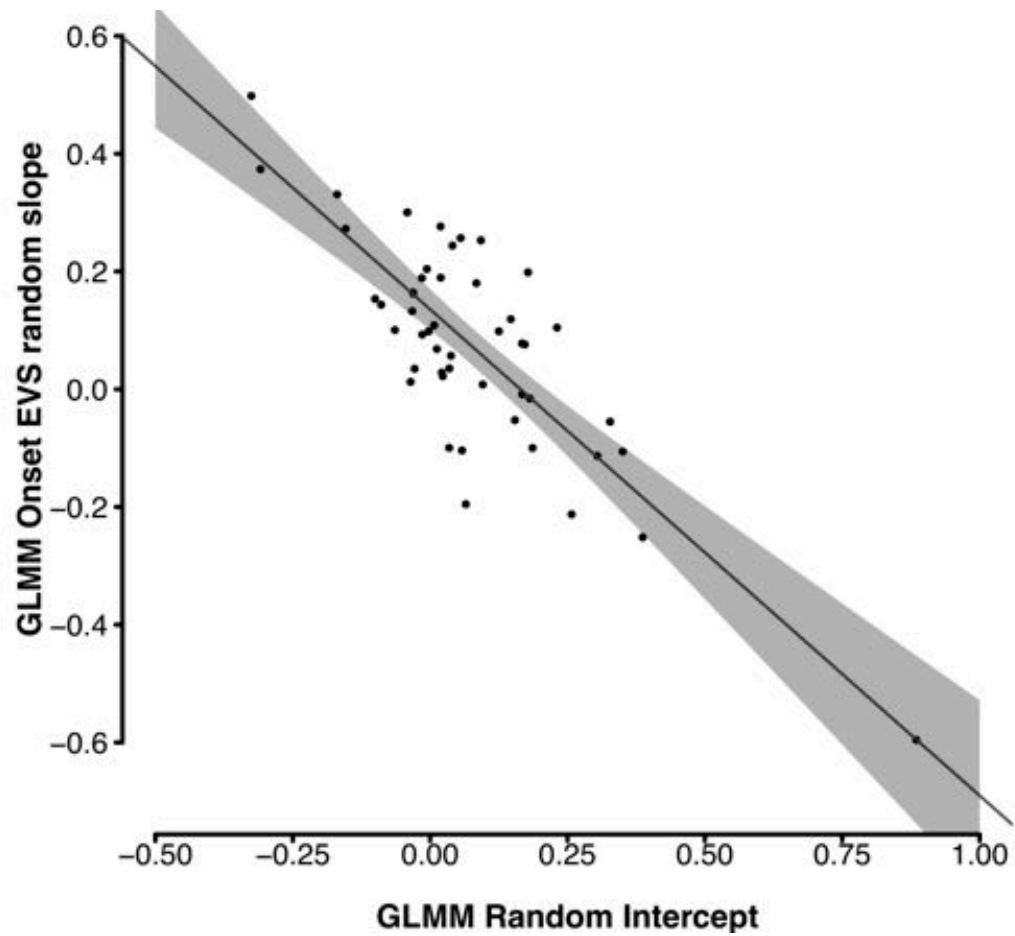
density distribution for EVS and FVS



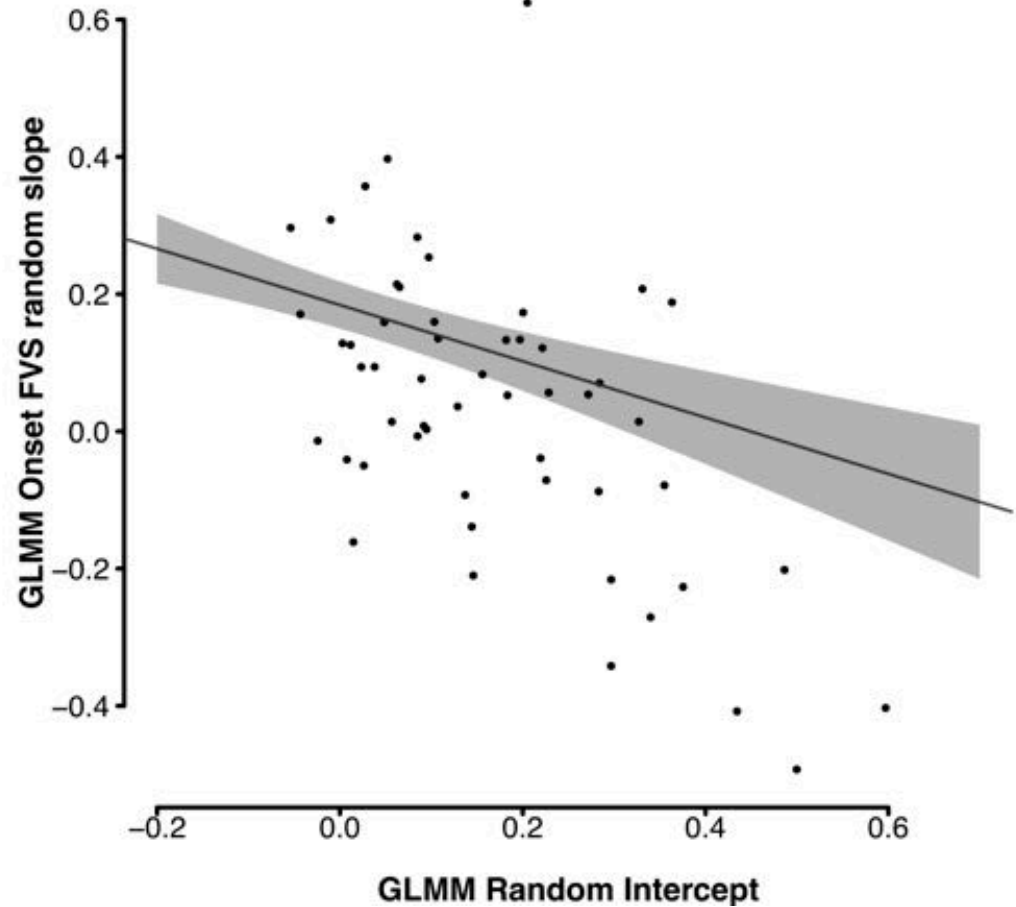
voice, finger and eye movements



the voice span affects tracking duration



**eye**



**finger**

fixation and tracking times as a function of articulation rate

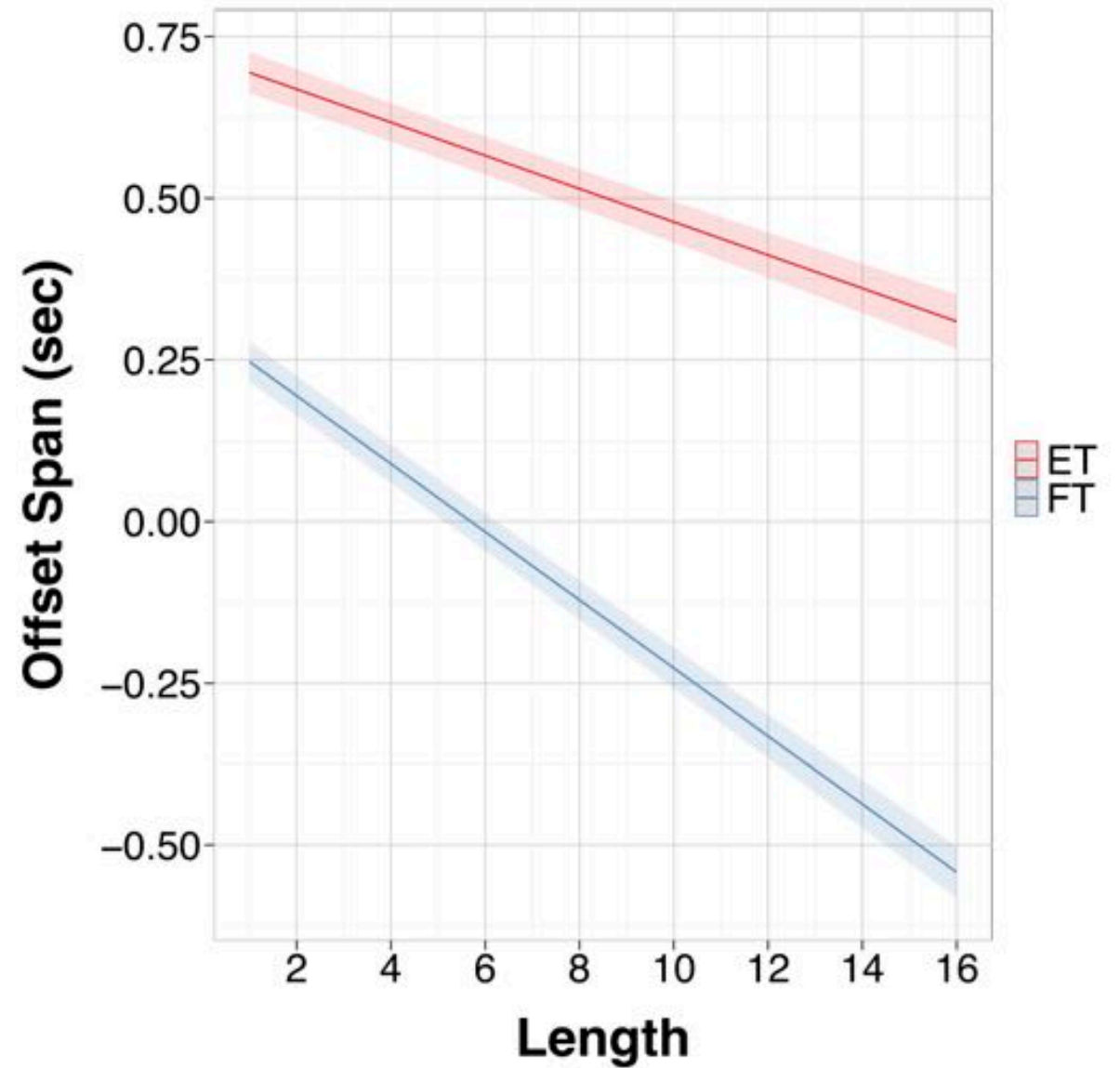
finger-pointing reading

what modulates  
the voice span?



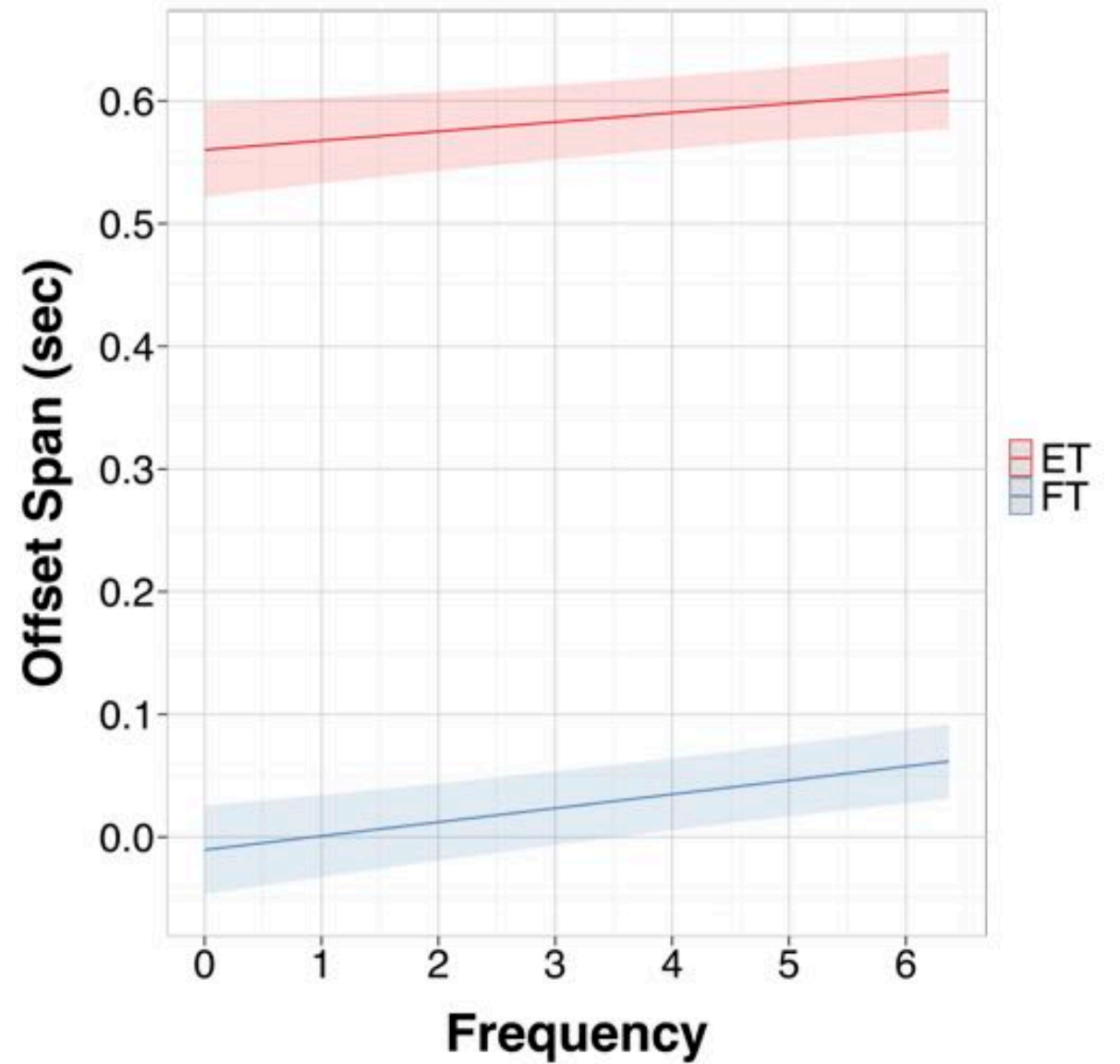
# buffer capacity

offset span ~  
(length + frequency + punctType) \* trackMode + (1| token) + (1| subject)

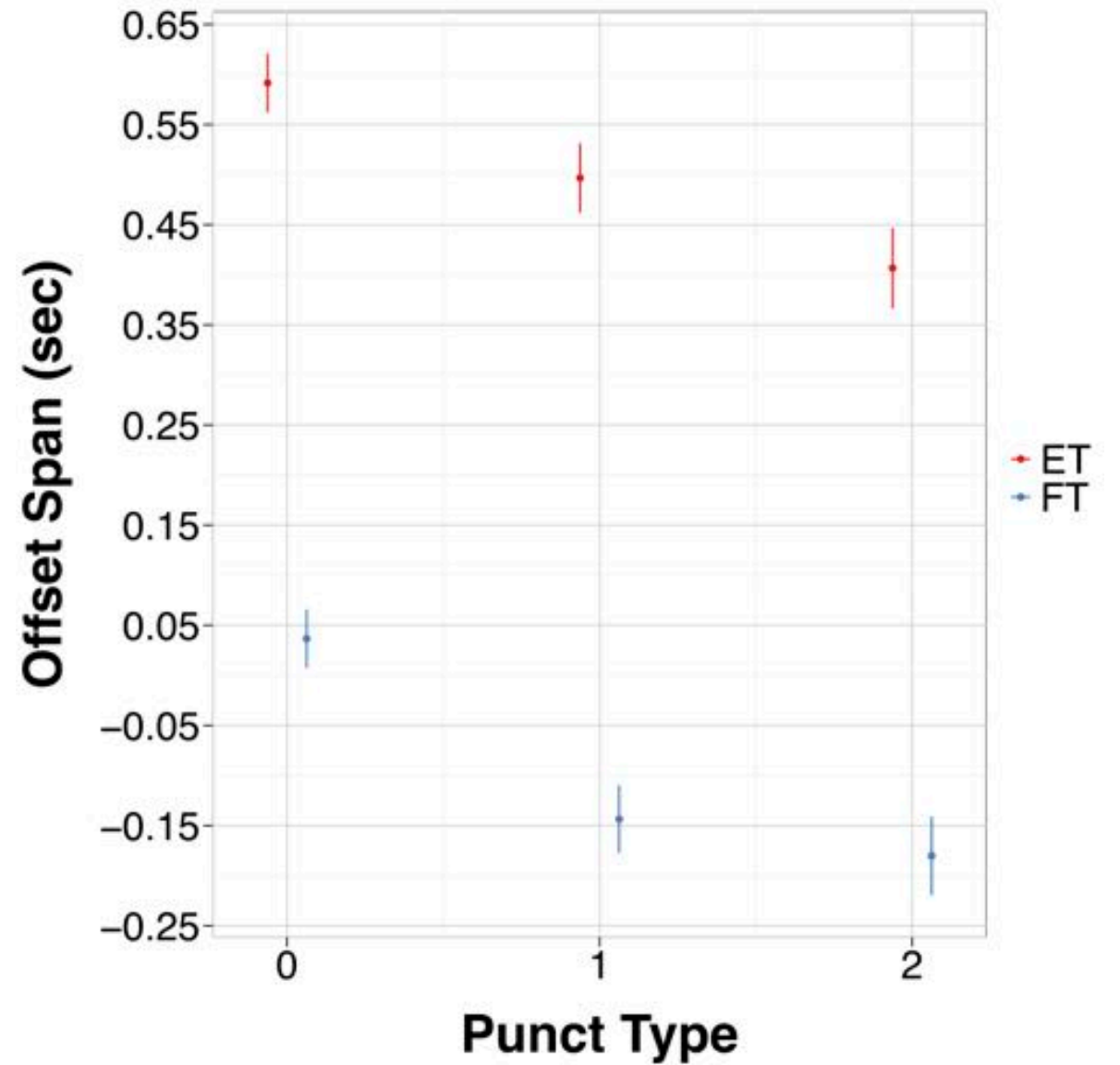




automaticity



implicit prosody



finger-pointing reading

what about  
multi-word chunks?

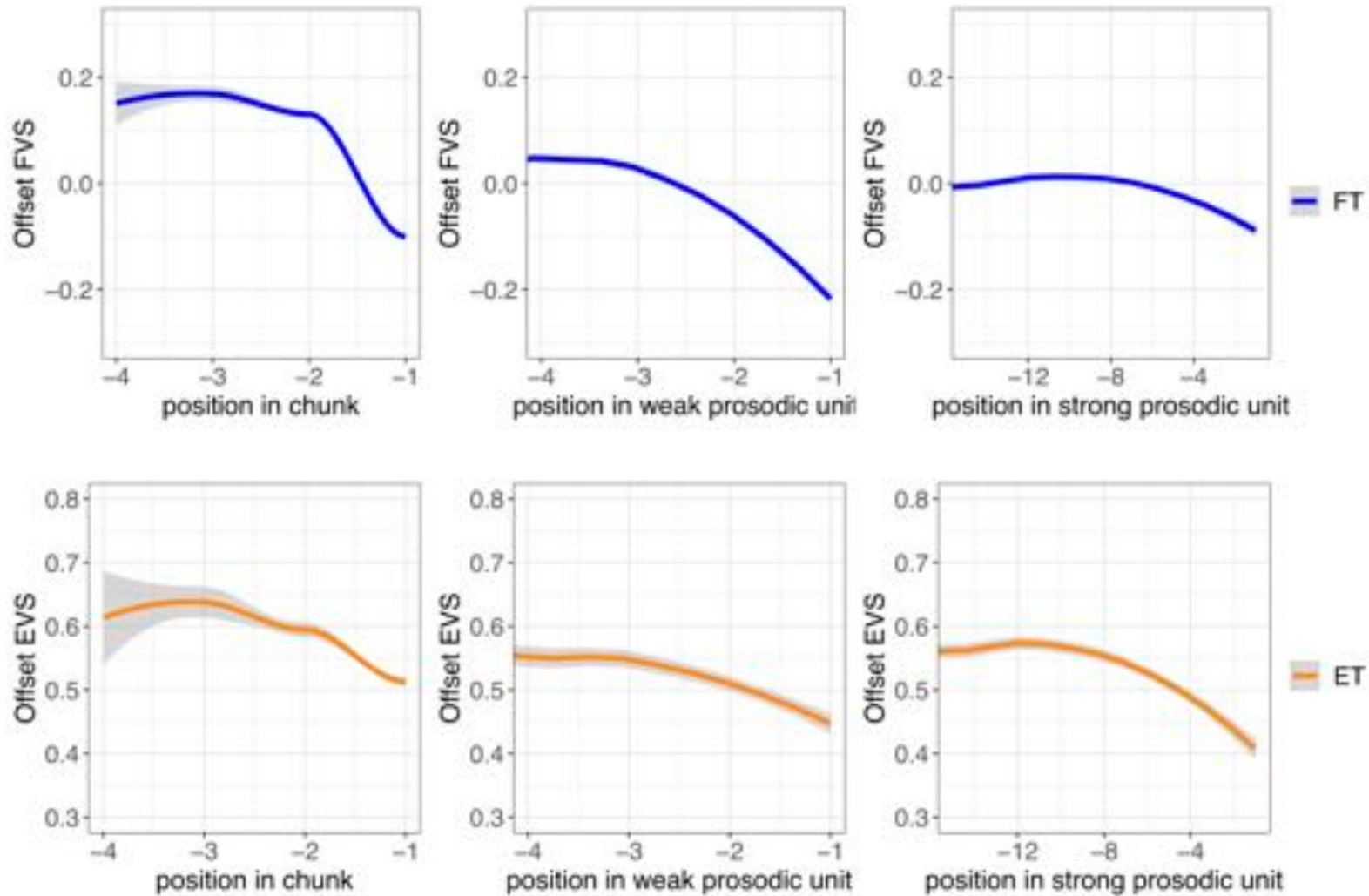


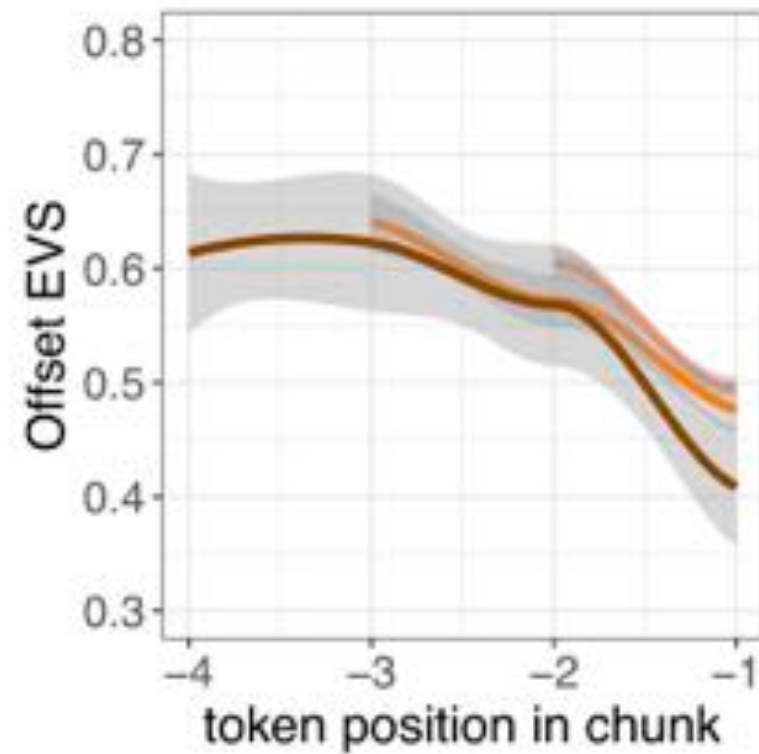
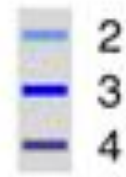
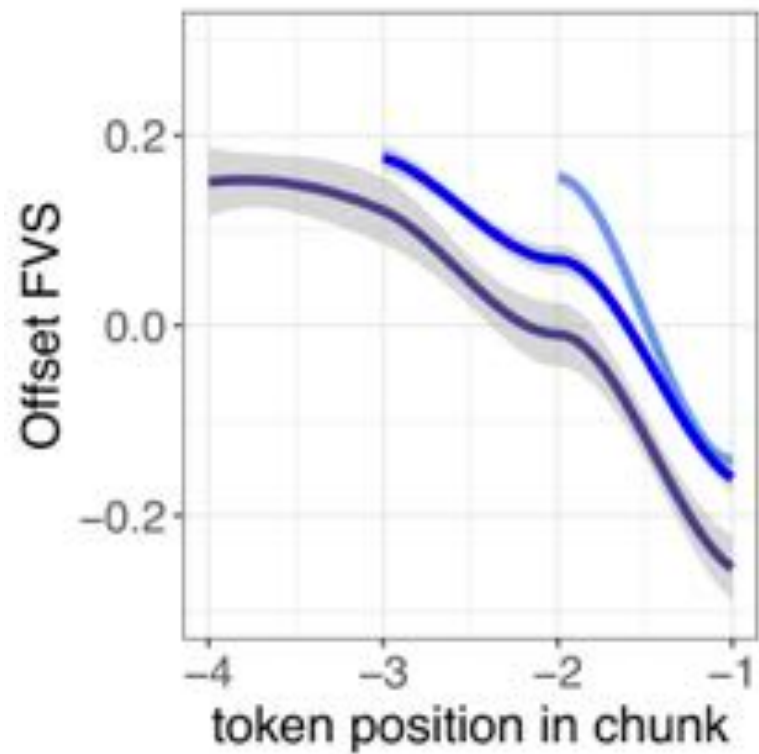
chunking: a  
non-recursive  
level of phrasal  
constituency

the black cat

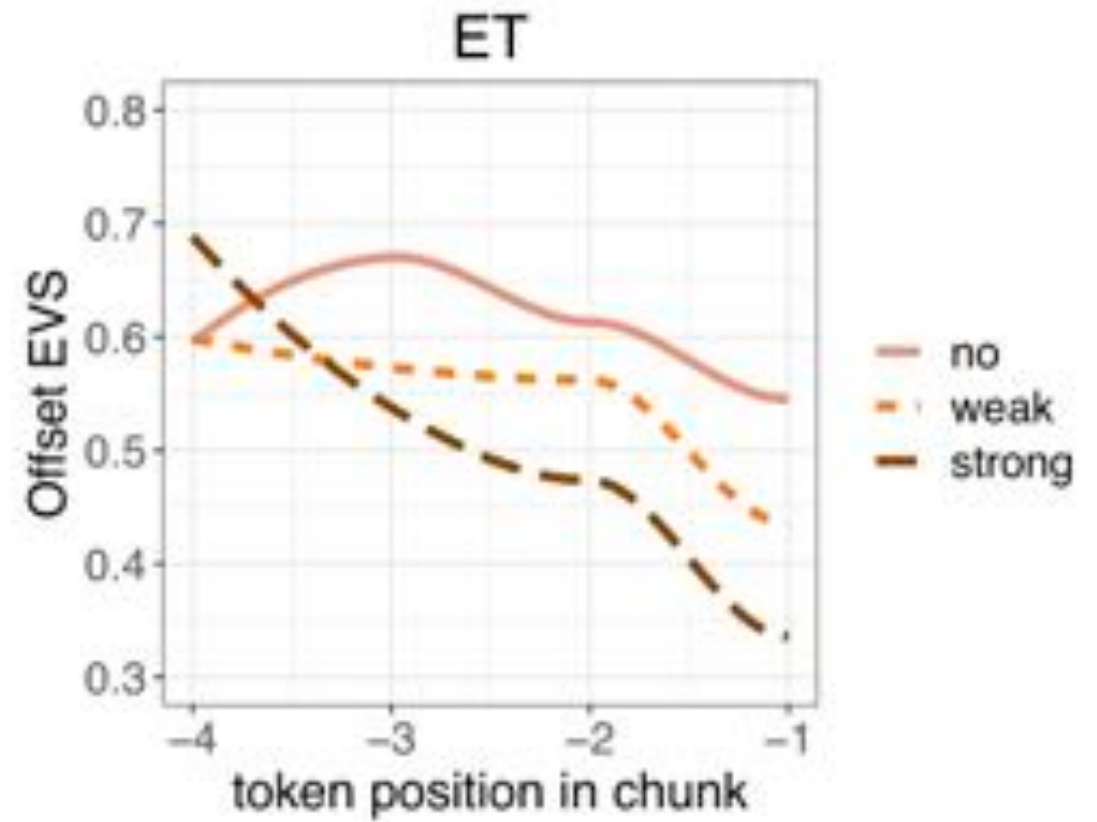
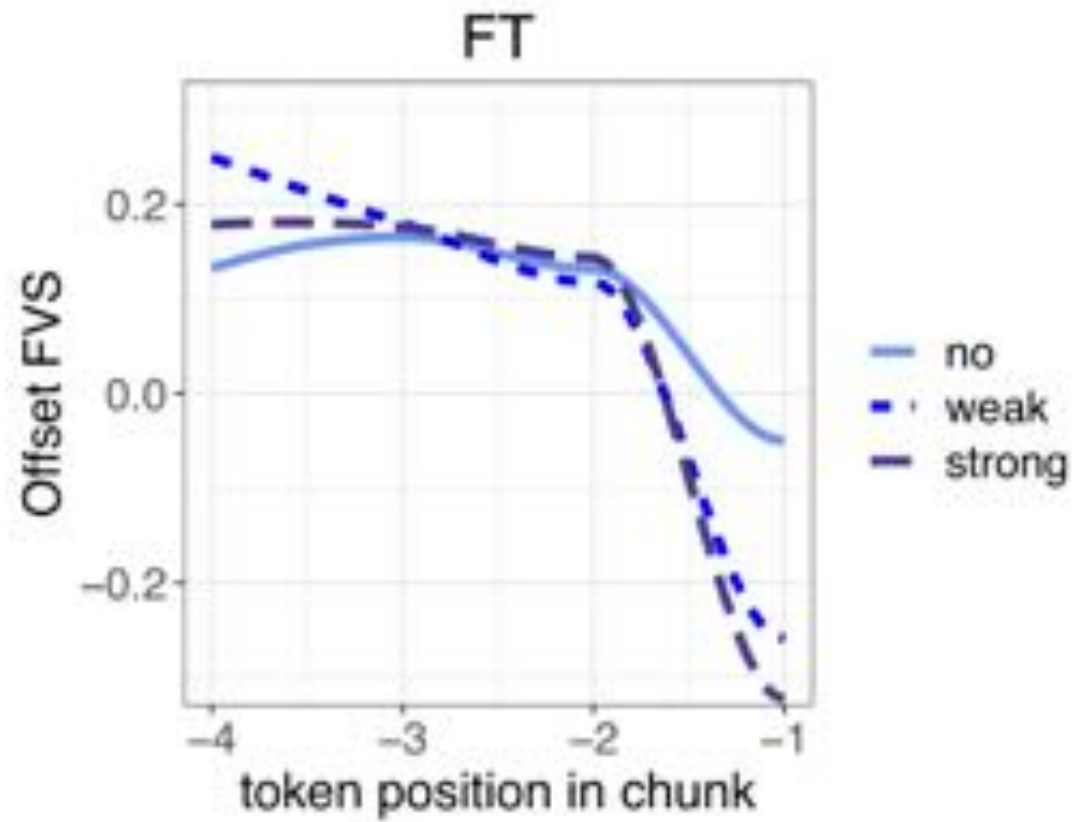
of my eldest daughter

# offset span at token position within chunk

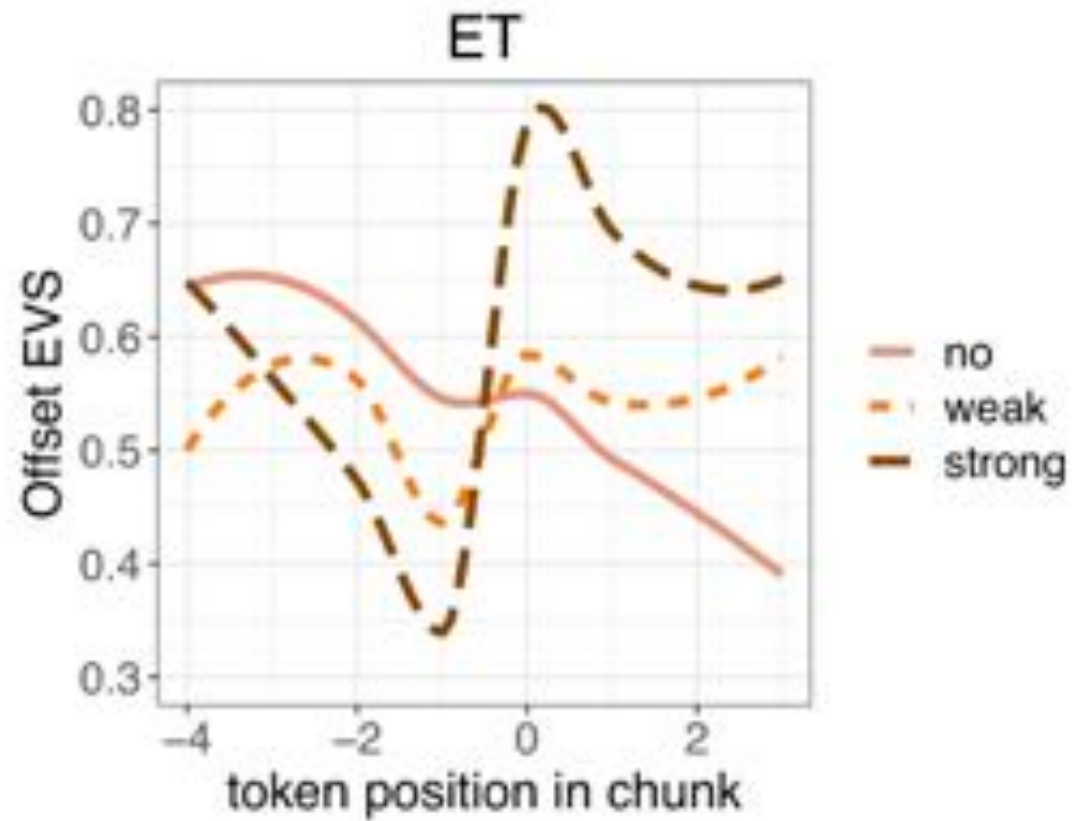
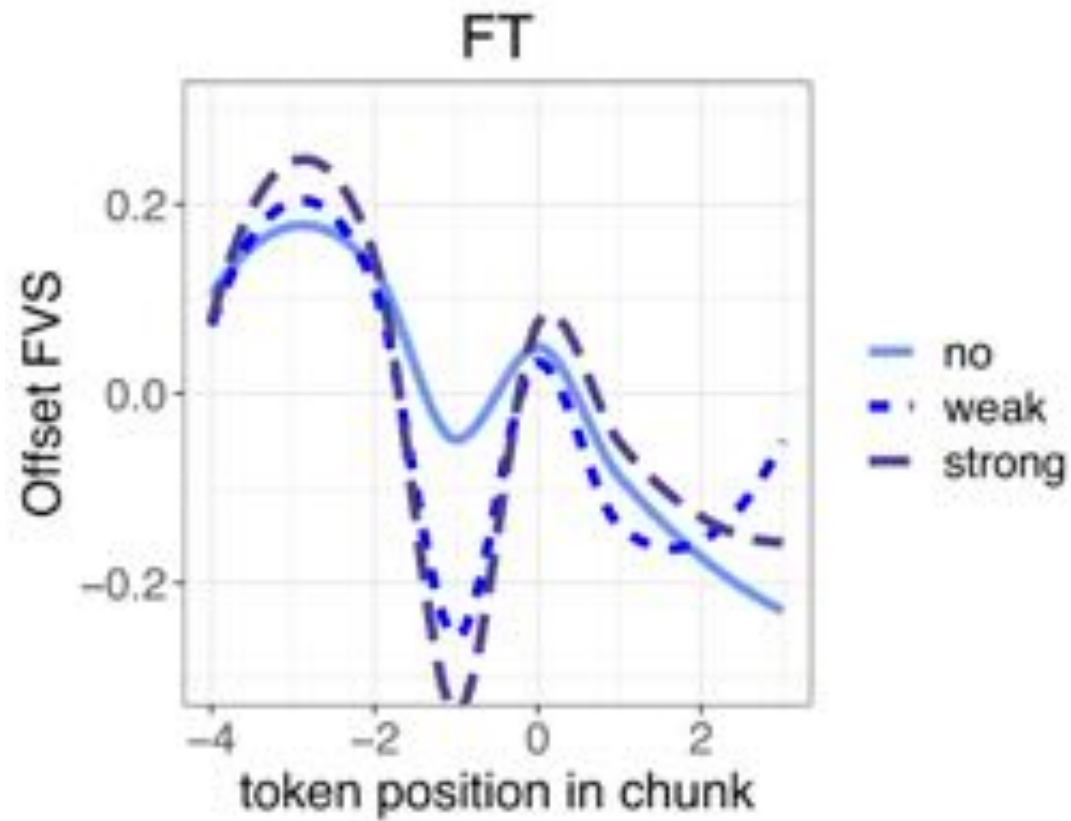




offset span for chunks of different length



prosodic effects



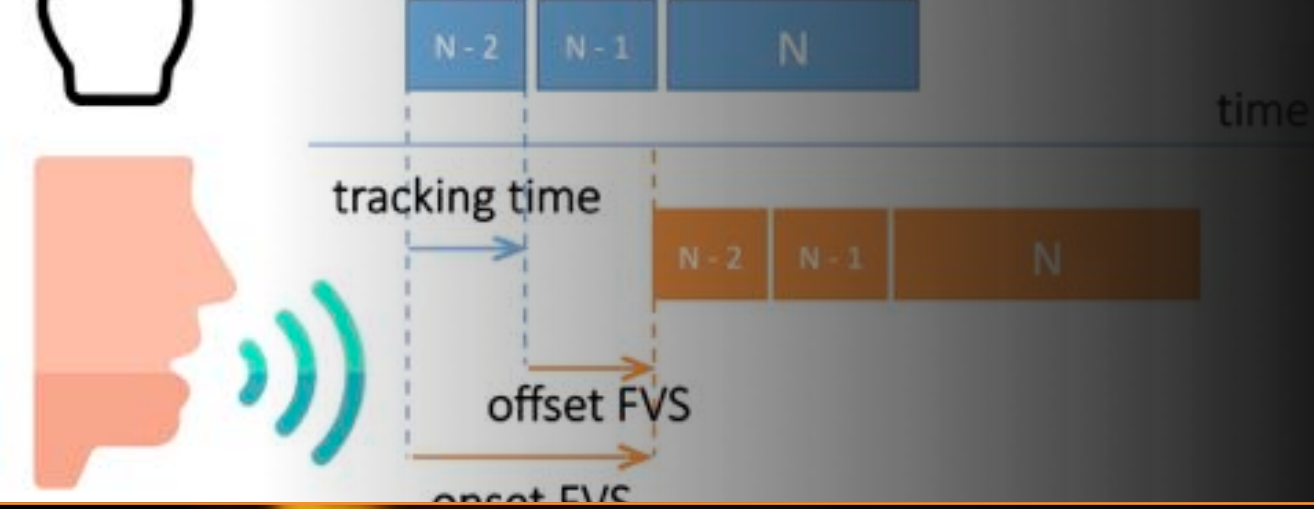
across boundary effects





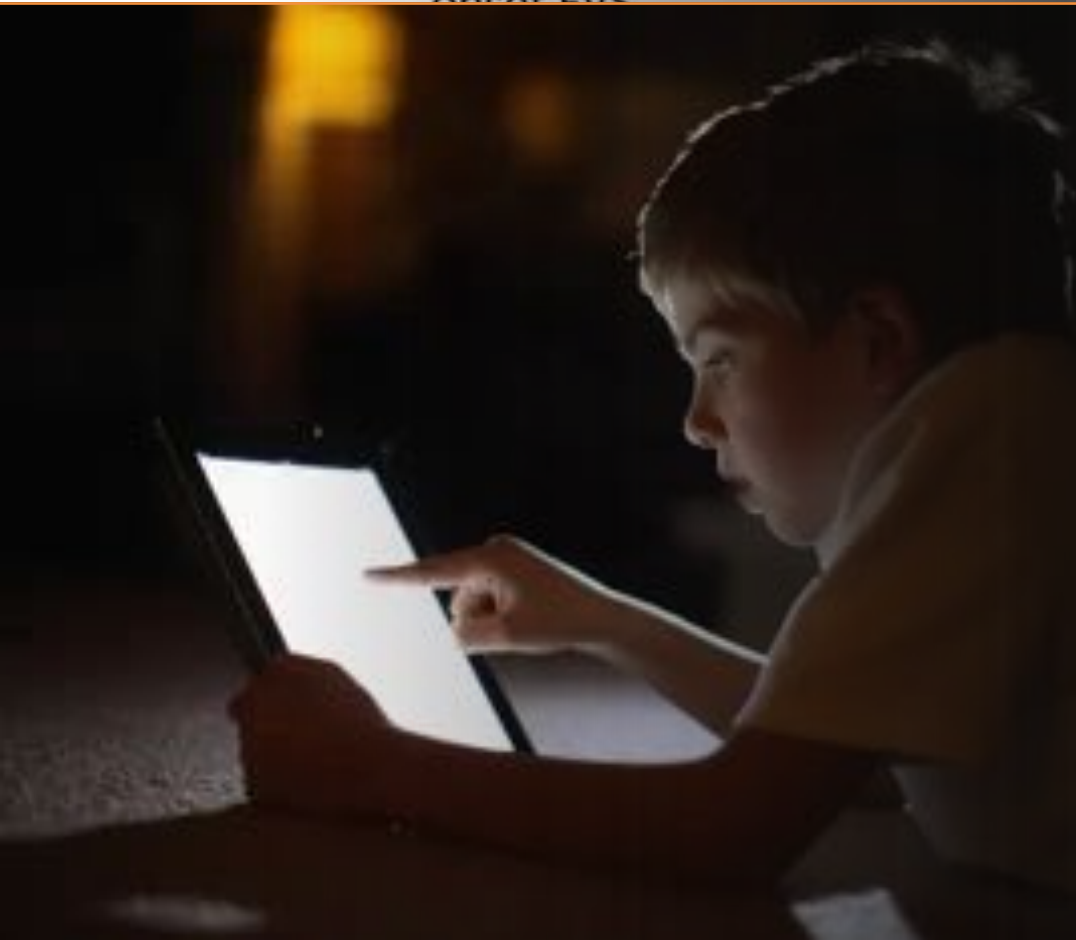
concluding remarks





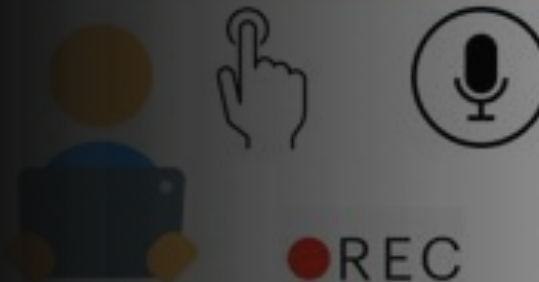
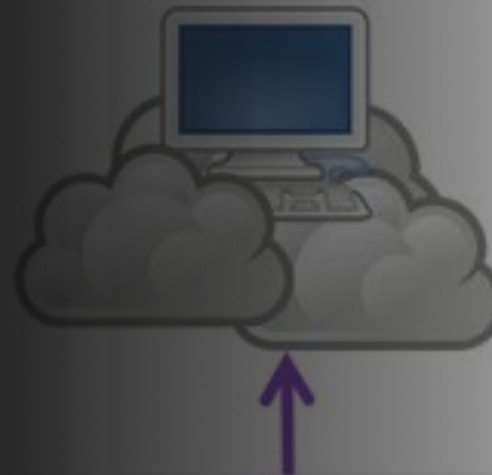
# voice spans & reading proficiency

- the voice span may vary across subjects depending on their **rate of articulation and memory buffer**
- spans vary even more prominently within the same reader, as they tend to **get larger** when the text contains **larger intonational and structural units**
- spans are modulated differently depending on tracking mode, with **the eye being mainly sensitive to larger text units**, and **the finger to smaller ones**
- all this evidence is indicative of an **optimally adaptive reading strategy**, whereby a proficient reader is able to adjust her articulation rate and phonological buffer to the structure complexity of a text



# simple & powerful

- relatively simple, noiseless and cheap, finger tracking offers ways of monitoring **reading development at scale**, using highly ecological reading tasks
- Its fine-grained information can profile reading skills at both the population and the individual level
- the time-bound, analogical signal of finger tracking makes it possible to observe **dynamic processing effects**
- we are currently experimenting the usage of our tablet in clinical contexts, where finger tracking combined with AI (speech processing) technology can offer unprecedented opportunities for reliable, efficient and noninvasive screening of **specific learning disorders**





thanks!

we gratefully acknowledge funding from PRIN grant no. 2017W8HFRX