Subtitled speech: Phenomenology of tickertape synesthesia

Fabien Hauw, Mohamed El Soudany, Laurent Cohen

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Exploratory Report

Subtitled speech: Phenomenology of tickertape synesthesia

Fabien Hauw [a,b,*], Mohamed El Soudany [a] and Laurent Cohen [a,b]

[a] Inserm U 1127, CNRS UMR 7225, Sorbonne Universités, Institut du Cerveau, ICM, Paris, France
[b] AP-HP, Hôpital de La Pitié Salpêtrière, Fédération de Neurologie, Paris, France

ARTICLE INFO

Article history:
Received 23 September 2022
Reviewed 27 October 2022
Revised 31 October 2022
Accepted 16 November 2022
Action editor Robert McIntosh
Published online 17 December 2022

Keywords:
Synesthesia
Tickertape synesthesia
Reading

ABSTRACT

With effort, most literate persons can conjure more or less vague visual mental images of the written form of words they are hearing, an ability afforded by the links between sounds, meaning, and letters. However, as first reported by Francis Galton, persons with ticker-tape synesthesia (TTS) automatically perceive in their mind's eye accurate and vivid images of the written form of all utterances which they are hearing. We propose that TTS results from an atypical setup of the brain reading system, with an increased top-down influence of phonology on orthography. As a first descriptive step towards a deeper understanding of TTS, we identified 26 persons with TTS. Participants had to answer to a questionnaire aiming to describe the phenomenology of TTS along multiple dimensions, including visual and temporal features, triggering stimuli, voluntary control, interference with language processing, etc. We also assessed the synesthetic percepts elicited experimentally by auditory stimuli such as non-speech sounds, pseudowords, and words with various types of correspondence between sounds and letters. We discuss the potential cerebral substrates of those features, argue that TTS may provide a unique window in the mechanisms of written language processing and acquisition, and propose an agenda for future research.

1. Introduction

Literate persons with the ability to form mental images, can voluntarily conjure visual images of words they are hearing, an ability afforded by the links between sounds, meaning, and letters which actually define literacy. Such images are vague and their generation requires substantial effort. However, as first reported by Francis Galton, “some few persons see mentally in print every word that is uttered (…) and they read them off usually as from a long imaginary strip of paper, such as is unwound from telegraphic instruments” (Galton, 1883). Galton noted that beyond this core subjective phenomenon, “the experiences differ in detail as to size and kind of type, color of paper, and so forth, but are always the same in the same person”. For instance, the botanist George

* *Corresponding author*. Inserm U 1127, CNRS UMR 7225, Sorbonne Universités, Institut du Cerveau, ICM, Paris, France.
E-mail address: fabien.hauw@orange.fr (F. Hauw).

https://doi.org/10.1016/j.cortex.2022.11.005

Henslow reported that the illusory words displayed “extreme restlessness” and that they “oscillated, rotated, and changed”. Galton also briefly discussed the involvement of attention, indicating that some individuals “attend to the visual equivalent and not to the sound of the words”, sometimes resulting in actual embarrassment, as reported by a fellow member of the Royal Society who told Galton that he “often craved for an absence of visual perceptions, they are so brilliant and persistent”.

Ever since this initial description, this phenomenon was subsumed under the broad label of synesthesias, a family of perceptual peculiarities, in which stimulation in one sensory or cognitive domain automatically induces an unusual and vivid perception in a distinct sensory or cognitive domain. Two instances of synesthesia which have received much attention are grapheme-color synesthesia, in which digits and letters, printed or even imagined, are perceived as colored, and sound–color synesthesia, in which the perception of a musical sound evokes the experience of color. The illusory perception of written words during speech perception was labeled “ticker tape synesthesia” (TTS), to which we will stick for consistency, although “subtitles synesthesia” would be a more readily understandable term.

Beyond being casually mentioned among the more than 100 types of synesthesia, TTS has received little scientific attention. Chun and Hupé (2013) estimated the prevalence of TTS at 7%, a likely overestimation due to sampling biases. They also noted that TTS is often associated with other types of synesthesia. Holm et al. (2015), over several samples of population, found 1.4% of “obligatory” tickertapers who experienced TTS unwillingly whenever hearing external speech or their own voice, or even when thinking in words. The rate rose to 8.7% for individuals in whom some orthographic imagery occurred only occasionally. Overall, a maximum incidence of a few percent would be in the range of other types of synesthesia (Johnson et al., 2013; Rothen & Meier, 2010; Simner et al., 2006). Although TTS has no significant impact on everyday life, Coltheart and Glick (1974) documented an advantage in spelling words backward, a task that may be facilitated by the use of a stable mental image of written words.

TTS stands out of the list of known synesthesias because, in synesthesia parlance, both the inducer (speech) and the concurrent (orthographic image) are language representations. We therefore propose that the potential interest of TTS goes far beyond the anecdote, as it may provide a unique window in the mechanisms of written language processing and acquisition.

In the brain of literate persons only, words are coded as series of letters, and dedicated processes allow for the translation from orthography to sound and meaning during reading, and for the opposite translation during spelling. Those novel features develop during the acquisition of reading, through delicate changes in visual and language areas of the brain, and changes in their connectivity (Dehaene et al., 2015; Dehaene-Lambertz et al., 2018). This complex tuning occurs under a variety of environmental and genetic constraints, resulting in substantial individual variability in eventual performance level. For instance, twin studies have shown significant heritability of reading abilities (Harlaar et al., 2007), while factors such as birth weight represent risk factors for reading disorders (Mascheretti et al., 2018). At one end of the performance spectrum stand the various types of developmental dyslexia (Friedmann & Coltheart, 2016), associated with atypical anatomical and functional brain features (Norton et al., 2015; Ramus et al., 2018). TTS should be understood, on this background, as a further form of atypical setup of the cerebral literacy system.

In a recent study we assessed the hypothesis, first put forward by Holm et al. (2015), that TTS would result from unusually intense top-down influences from phonological and lexical representations, embodied in left perisylvian language areas, onto orthographic codes subtended by the left ventral occipitotemporal cortex (Hauw et al., 2022). Using fMRI in a single tickertaper, we showed that a set of left-hemispheric areas were more active in the tickertaper than in controls during the perception of normal than reversed speech, including perisylvian areas involved in speech processing, and the Visual Word Form Area, a left occipitotemporal region subtending orthography. Those areas were identical to those involved in reading, supporting the construal of TTS as “upended reading”. Using dynamic causal modeling, we further showed that TTS indeed relied on increased top-down flow of information.

Here, as a first step towards a broader understanding of TTS, we endeavored to draw a fuller portrait of its core features and of its variations along the multiple dimensions conjectured by Galton, including visual and temporal features, triggering stimuli, voluntary control, interference with language processing, etc. We do so based on an extensive questionnaire, and on the description of synesthetic percepts elicited by auditory stimuli such as non-speech sounds, pseudowords, and words with various types of correspondence between sounds and letters. We discuss the potential cerebral substrates of those features, and propose an agenda for future research.

2. Methods

2.1. *Recruitment of participants*

We created a short description of what TTS is, and recruited participants by broadcasting it through email and social networks, targeting particularly groups of students and university members, and groups devoted to synesthesia, psychology, and neuroscience. Among individuals who answered us, we included those who reported having TTS, were over 18 years old, and had no neurologic or psychiatric history. Participants provided written informed consent and the study was approved by the institutional review board of the INSERM (protocol C13-41). We proposed to all participants (i) a questionnaire mostly aiming at characterizing the subjective features of TTS, and (ii) a series of auditory stimuli for which they were asked to report the content of the elicited synesthesia. No objective tests were available to ensure that all self-reported synesthetes actually had TTS, contrary to other synesthesias in which it is at least possible to assess the consistency of induced-concurrent associations, for example in grapheme-color synesthesia (Carmichael et al., 2015).

2.2. *Questionnaire on ticker tape synesthesia*

The full content of the questionnaire is reported as Supplementary material. The survey addressed the following topics.

2.2.1. *Demographics*

Participants were asked about their gender, age, profession, native language, education level, reading proficiency, handedness, potential auditory or visual impairments, learning disabilities. We asked them to report their associated synesthesias if any, including: space-time (the experience of time as a spatial construct), sound–color (involuntary experience of colors when hearing sounds), number-space (ordinal sequences such as numbers map to stable spatial locations in mental space), grapheme-color (associations between letter and colors), digit-color (associations between digits and colors), ordinal-linguistic (ordinal sequences such as numbers or days, have personalities, emotions or gender), auditive-gustatory (experience of tastes that are triggered by the perception of non-taste related words), and perfect pitch. They were also asked about TTS and other synesthesias among family members.

2.2.2. *Subjective features of TTS*

We investigated the following dimensions of TTS experience (see Supplementary material for full details).

Onset and awareness of TTS: TTS being a developmental phenomenon, the first questions aimed at determining when the synesthesia emerged, and when and how participants realized it was uncommon.

Stimuli that trigger and modulate synesthesia: On a rough definition, TTS is triggered by speech. In order to determine the detailed features of speech actually triggering and modulating TTS, we asked the following questions. Is synesthesia triggered by external speech, by one's own speech, etc.? Do auditory parameters such as emotional content or gender of the speaker modulate TTS perception? What extent of voluntary control do participants have over the phenomenon?

Visual features: The very substance of TTS being mental visual images of unusual vividness, we asked the following to characterize the features of images as precisely as possible. Do participants perceive the subtitles within their own mind or projected in the external world? Are subtitles always perceived in the same place? Are they written in upper- or lower-cases, are they colored, etc.? How many words are perceived at the same time? How are number words perceived?

Temporal features: Although actual mental chronometry may be out of the reach of introspective reports, we tried to identify the gross relative temporal unfolding of the auditory and visual components of TTS, asking the following questions. Is there any delay between speech perception and the onset of synesthesia, and do subtitles persist for some time after speech stops?

Non-word triggers of TTS: To the above exploration of the main triggers of TTS, we added a set of questions addressing stimuli at the edge between speech and non-speech. Is TTS triggered by utterances in a foreign language, by novel words, by different classes of non-speech sounds such as sneezing, meowing, or the sound of a bell?

Interference with daily life: Finally, in an attempt at identifying objective markers of TTS, we asked whether participants found any advantage or disadvantage to TTS in their daily life. For instance, we tried to determine what happens when reading a book near someone speaking, or when watching a movie with subtitles.

2.3. *Assessment of TTS with auditory stimuli*

Participants were presented with auditory stimuli, and were asked to write down whatever subtitles they may induce. They could replay each sound as often as wished. We used the following types of sounds.

2.3.1. *Non-speech stimuli*

We selected audio files from an open sound library (www.lasonotheque.org), including (1) animals sounds (meowing, bleating, barking and crowing), (2) human noises (laugh, cough, yawning and sneezing), (3) sounds of inanimate objects (car engine, ring tone, iron bar falling, gunshot), and (4) short wordless musical tunes.

2.3.2. *Words with exception spelling*

Exception words do not obey usual phoneme-to-grapheme correspondences, and spelling them out therefore requires access to stored lexical knowledge of orthography (e.g., YACHT or OIGNON in French). By presenting TTS individuals with exception words, we wished to determine whether such lexical knowledge was reflected in the content of subtitles. We selected 10 French exception words that had at least 3 phonemes, and generated the corresponding audios files using the ttsreader.com speech generation software.

2.3.3. *Homophone words*

Homophone words share the same pronunciation but differ in meaning. When they also differ in orthography they are called non-homographs (e.g. LEAK vs LEEK, or CONTE vs COMPTE in French). By presenting such words in spoken form, we wished to determine whether subtitles featured only a single spelling, possibly the most frequent one, or several concurrent spellings. We selected ten auditory word forms, based on 2 criteria: each should correspond to at least 3 words with different spellings, and one of those spelling should have a substantially higher frequency of occurrence than the others, according to the to the lexique.org database. We generated the corresponding audio files as before. In order to analyze how synesthetic spellings were selected among the possible spellings, we computed the percentage of participants who reported each spelling over the 10 stimuli, and modelized those values with a general linear mixed-effects model (GLMM), with log frequency and grammatical category (noun vs other) as fixed factors, and stimulus identity as random factor.

2.3.4. *Pseudowords*

Pseudowords are well-formed phonological strings which differ from real words only in that they have not been encountered before and are therefore absent from the mental lexicon (e.g. CHADOURNE in French). They can be spelled-out only on the basis of the statistics of phoneme-to-grapheme correspondences. By presenting pseudowords, we wished to determine whether such non-lexical sound-to-spelling mechanisms also controlled the content of subtitles. Using the UniPseudo tool available on http://www.lexique.org we created 10 6-phonemes pseudowords, and generated the corresponding audio files as previously described.

3. Results

We identified 26 participants with TTS, and asked them to answer a detailed questionnaire, including mostly retrospective questions on the synesthesia, plus a set of auditory files for which they were asked to transcribe the triggered subtitles. We will report in turn the participants' general characteristics, the individual history of TTS, the subjective visual and temporal features of the subtitles, the specifics of subtitles evoked by number words and other types of auditory stimuli, the varieties of overt and covert speech triggering TTS, and finally the participants' ability to exert voluntary control over TTS synesthesia.

3.1. *Demographic features of participants with TTS*

We first determined the demographic features of participants, and their personal and family history of synesthesia.

Demographics: All 26 participants were French native speakers (see Table 1; 9 male; median age: 40 years). Birth was normal in 18 participants (69%), 0 were born prematurely, and 8 had other birth complications. Most of them (25/26, 96%) had a higher educational level, while one stopped after middle school. Twenty-four (92%) had normal sight with or without correction. All had normal hearing.

Cognitive development: All participants had normal cognitive development. None reported dyspraxia, autism spectrum disorder, dyslexia or delayed language acquisition.

Reading habits: Regarding reading habits during childhood, 20 participants (77%) considered themselves as avid readers, 4 (15%) as standard readers, and 2 (8%) as occasional readers. During adulthood, 17 participants (65%) still read a lot, 6 (23%) were standard readers, and 3 (11%) were occasional readers.

Associated types of synesthesia: Eighteen participants (69%) reported other associated types of synesthesia (Table 2), mostly space-time in 12 (46%), number-space in 9 (35%), and sound–color in 7 (27%). Ten participants (38%) had more than one associated synesthesia, and 1 participant (4%) had 7 associated synesthesias.

Family history: Nine participants (35%) had other TTS subjects in their family, 2 (8%) did not, and 15 (58%) did not know. Note that 3 participants belonged to one family, and 2 others to another family. Moreover, 7 participants (27%) had members of their family with other types of synesthesia, 2 (8%) did not, and 17 (65%) did not know.

3.2. *History and general features of TTS*

We then analyzed at what age TTS emerged, and how participants became aware of this peculiarity.

Age of onset: Reading acquisition is a logical prerequisite to the emergence of TTS. Indeed, nineteen participants (73%) reported that TTS had been present since reading acquisition ($n = 11$) or for as long as they could remember ($n = 8$); 3 participants reported a later onset during adolescence; the remaining 4 did not know.

Awareness of TTS: Most participants (19/26, 73%) discovered that TTS was an unusual phenomenon only during adulthood (>20 y.o.), 2 (8%) during adolescence (16–20 y.o.), and 3 (12%) during childhood (<16 y.o.). This discovery was fortuitous in 8 participants (31%), 6 participants (23%) became aware of this peculiarity through the advertisement for the current study, and 4 participants (15%) through reading a journal article or watching a documentary. Eight participants did not answer. Finally, 19 participants (73%) had mentioned TTS to their family or friends, 2 had been deliberately hiding it, while the remaining 5 (19%) had simply never mentioned it, not knowing that it was an unusual phenomenon.

An advantage or a disadvantage?: Only 4 participants considered that TTS had neither a positive nor a negative impact on their daily life activities. For 12 participants (46%) it was both an advantage and a hindrance, only an advantage for 6 participants (23%), and only a hindrance for 4 (15%). The main subjective advantage was the support for spelling words correctly, and the main disadvantage was difficulty focusing their attention toward a single speaker in crowded places.

Attention and control: For 18 participants (70%), TTS was overall an automatic phenomenon on which they had little or no control. Still, when listening to someone, 7 participants (27%) could voluntarily prevent TTS to occur, while only 5 (19%) could do so when they were speaking themselves. When participants did not pay attention to external speech, e.g. in a

Table 1 – General characteristics of participants.

	n = 26
Male	9 (35%)
Median age (range)	40 (18–70)
Problems at birth	8 (31%)
Higher education	25 (96%)
Normal cognitive development	22 (85%)
Family history of TTS	9 (35%)
Family history of other types of synesthesia	7 (27%)
Associator synesthetes	18 (69%)
Projector synesthetes	2 (8%)

TTS: ticker-tape synesthesia.

Table 2 – Associated synesthesias.

	n = 26
Space-time	12 (46%)
Number-space	9 (35%)
Sound–color	7 (27%)
Grapheme-color	6 (23%)
Digit-color	6 (23%)
Ordinal-linguistic	6 (23%)
Perfect pitch	3 (12%)
Auditive-gustatory	2 (8%)
None	8 (31%)

restaurant, the unattended speech still triggered TTS in 15 participants (58%).

3.3. Speech modalities triggering TTS

Participants were then asked to specify which features of speech input were required to trigger TTS.

Types of speech input: The main triggers of each participant are reported in Fig. 1A. In all participants, TTS occurred when watching someone else speaking (26/26, 100%). In most of them (22/26, 85%), TTS persisted even when the speaker was out of sight. In most participants, their own overt speech (23/26, 88%) and covert speech (23/26, 88%) generally triggered TTS. Moreover, TTS was generally triggered by musical lyrics and movies, by covert reading in half the participants, and a few participants reported seeing subtitles while dreaming.

Non-speech sounds: Other types of sounds triggering TTS according to the questionnaire are reported in Fig. 1B. TTS occurred in 25 participants (97%) when listening to novel words (e.g. a brand name), among who 5 reported a different TTS from usual (e.g. two reported "blurry words"); one did not know. TTS occurred in 12 (46%) participants when listening to human noises (e.g. laughs or cries), 11 (42%) for non-human noises (e.g. bird song), and in only 5 (19%) when listening to music without lyrics.

Voice features: Finally, some speech variations influenced the visual properties of subtitles. Subtitles were modified when hearing emotional speech in 9 participants (35%; e.g. "it may change the color of words", "letters may be blurred or shaking when I'm moved", "words are larger for strong emotions"), when hearing louder vs. softer voice in 7 (27%; e.g. "words are bigger for louder voices"), when hearing song lyrics in 5 participants (19%; e.g. "words are less visible than in normal speech", "words seem to undulate according to the melody"), while pitch had an impact only in 1 participant. The gender of the speaker had no reported impact. In 10 participants (38%), none of these factors affected the synesthesia.

3.4. Visual features of subtitles

We then collected information on the subjective visual features of subtitles evoked during speech perception.

Projection or association: Grapheme-color synesthetes are commonly distinguished in "projectors" and "associators", as defined by Dixon et al. (2004) depending on their external or internal perception of letter colors. In the case of TTS, 18 participants (69%), visualized subtitles within their internal mental space, while in 2 participants (18%), subtitles were perceived as projected in the outside world. Six participants (23%) reported variability between these two options, and one participant did not answer.

Location: The localization of subtitles was definite and stable in 18 participants (69%). Among those, subtitles were near the speaker's mouth in 4 (15%), or for instance "roughly at the same place as film subtitles, with a right or left shift depending on the origin of the voice", "in the center of my visual field", "behind my eyes". The location was definite but variable in 4 participants (15%). Finally, 4 participants (15%) reported no well-defined location.

Distance: Subtitles were considered as "nearby" in 15 (58%) participants, "far away" in 2 participants, 2 others chose both options, while 4 (15%) provided a precise location specified either in terms of distance (50 cm and 3 m) or more specifically (close to the speaker or within her own head). Three did not know.

Visual attributes: We then explored the font and color of the subjective display. Nineteen participants (73%) variably visualized words in lower- or upper-case, while 7 (27%) only in lower-case. No participant visualized words only in upper-case. Twenty participants (77%) visualized words written only in black, white or gray, while 4 others (15%), who all had associated grapheme-color synesthesia, saw colored letters (some including black/white letters). Two participants did not know. Character size was larger than in regular books in 12 participants (46%), had the same size in 8 (30%), and was smaller in one. Five participants did not know.

Content of subtitles: Seventeen participants (65%) perceived several words at the same time, 6 (23%) were limited to one word at a time. The span was reported as variable by 2 participants (8%). One participant could not answer.

Interaction with reading: When reading in a silent place, 12 participants (46%) visualized only the printed words, 12 (46%) perceived the printed text itself, plus a synesthetic version of the same text. The pattern was variable in 2 participants.

When reading while surrounded by people talking, TTS interfered with reading in 23 participants (88%). In 16 participants (62%), this resulted in an mingling of printed words and subtitles (e.g. "speech is intruding in the written text" or "polluting" it), 4 participants (15%) visualized both texts but in distinct mental spaces (e.g. "I'm seeing the subtitles 'above' the printed text", or "I can arrange the two channels in two parts of my head"), one managed to visualize only the printed words when focusing his attention, and 5 (19%) did not know.

Subtitled movies: When watching movies with subtitles, 11 participants (42%) saw both the real subtitles, and synesthetic subtitles in the original language spoken by the actors, sometimes at different locations. Some participants reported that for this reason watching subtitled movies required massive attentional effort. Nine participants (34%) saw only the real subtitles. In 4 participants (15%), the pattern was variable and depended on the familiarity of the spoken language.

3.5. Temporal features of subtitles

Next, we asked participants to describe how the content of TTS was dynamically updated during continuous speech perception.

Delay between speech and TTS: The onset of the synesthesia was immediate in 19 participants (73%), delayed in 6 (23%). One participant did not know.

Updating of subtitles: The content of TTS was updated by chunks of several words appearing simultaneously in 8 participants (31%), one word at a time in 7 (27%), only letter by letter in 1, and in a variable manner in 10 (38%). New words appeared at the same location as those they were replacing in 14 participants (54%). They scrolled from one side to the other "as if on a conveyor belt" in 9 (35%) participants, while 2 others

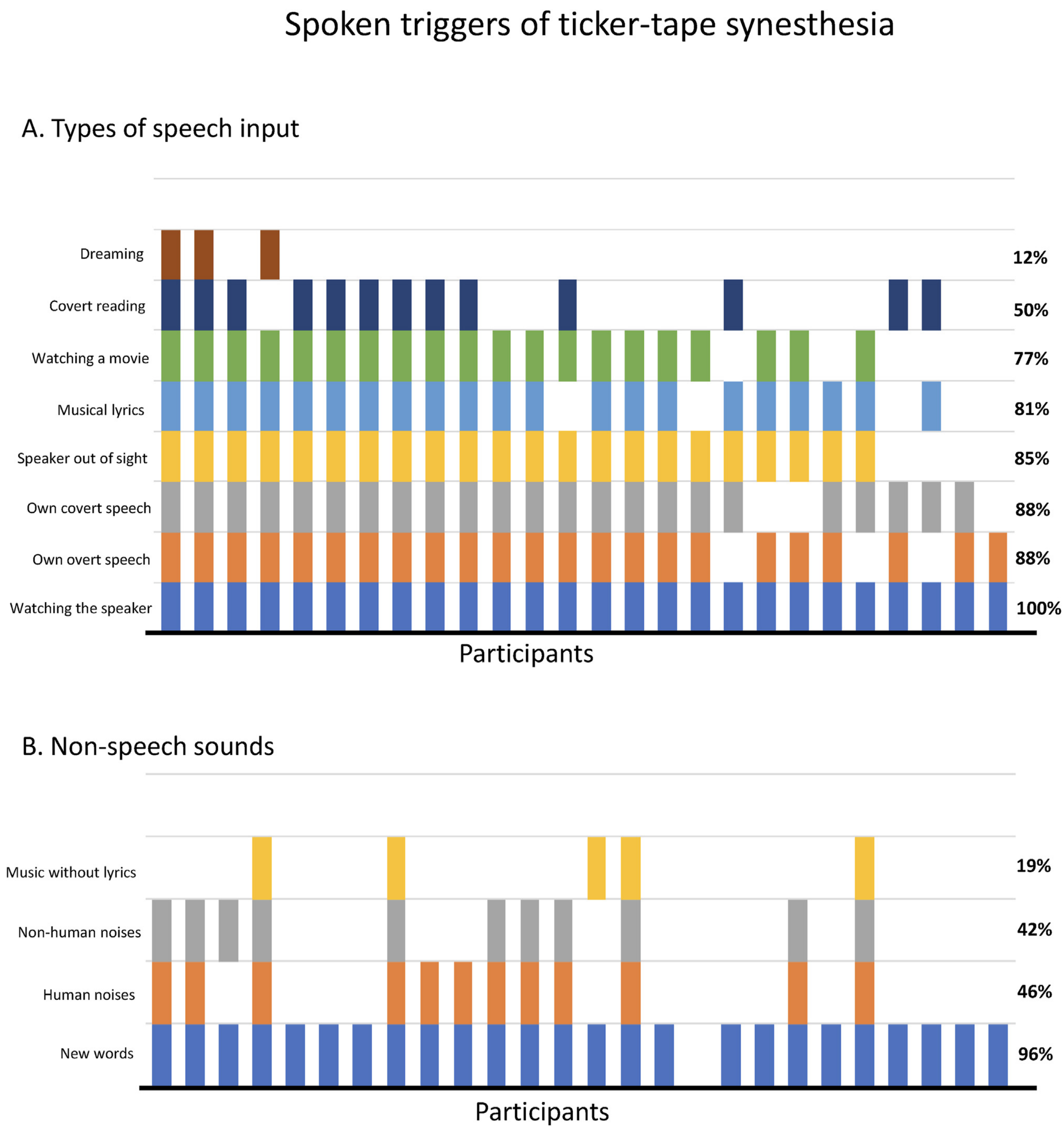

Fig. 1 — A. Summary of stimuli triggering ticker-tape synesthesia (TTS) based on the 26 participants' introspection. Each vertical bar represents a participant. Participants are sorted by decreasing number of trigger types. Each horizontal color line represents one trigger type. Triggers are sorted by decreasing number of concerned participants. Speech emitted by a visible speaker was a constant trigger, other varieties of speech triggered TTS in at least 20/26 participants. Covert reading was reported as a trigger in only 13/26 participants. TTS during dreaming was reported by only 3 participants. B. Summary of non-speech sounds triggering TTS based on the participants' introspection. The participants are sorted in the same order as in A. Each horizontal color line represents one trigger type. Triggers are sorted by decreasing number of concerned participants. Almost all participants reported TTS when listening to novel words, while only 5/26 (19%) when listening to music without lyrics. Percent values on the right represent the proportion of participants in whom TTS was elicited by the corresponding trigger.

referred to "bubbles" or "a prompter". Two participants did not know.

Persistence of subtitles: After their appearance, subtitles persisted briefly in 17 (65%) participants, disappeared immediately in 3 (12%) and had a more variable duration in 5 (19%). Complementing this admittedly vague labeling, a participant mentioned that the last word of an utterance lasted longer as it was not replaced by the following one, and another one that the duration of the subtitle was shorter during fast speech and longer if the speech flow was slower. Those anecdotal reports suggest that subtitles have a perceptible intrinsic duration, which was shortened by the occurrence of subsequent words. One participant did not know.

3.6. *Synesthesia for number words*

Number words and ordinary words all feature phonological, orthographic, and semantic representations. However, in addition to those shared formats, numbers may also be represented as Arabic numerals, requiring dedicated mechanisms for recognizing numerals and transcoding them to and from other word representations. We therefore explored to what extent number words would elicit alphabetic (e.g. "thirty") or Arabic (e.g. "30") subtitles, depending on the context in which number words were used.

Numbers denoting quantities: When numbers were used to denote quantities (e.g. "there are 35 people in the room"), 14 participants (54%) visualized them as Arabic digits, 5 participants (19%) as a mix of digits and letters, and 3 participants in alphabetic form only. Four participants (15%) did not know.

Indefinite article: In French, the same word "un/une" is used to express the indefinite article "a" and the unity numeral "one". When the number was used as an indefinite article (e.g. "j'ai UN ami"), 23 participants (88%) saw it in letters, 1 in digits, one saw both letters and digits, and one did not know.

Mental calculation: When performing mental calculation, 24 participants (92%) mentally saw numbers as Arabic digits, one saw both digits and letters, and one did not know. Moreover, one participant described visualizing geometrical shapes which combined depending on the operation, and another perceived numbers as dots on dice.

Date and hour: When hearing number words in the context of dates, 20 participants (77%) visualized them as digits, 4 saw a combination of digits and letters, and 2 did not know. When numbers were used to express time, 16 participants (62%) saw digits, 3 (12%) saw letters (accompanied by a clock face in one case), 4 (15%) saw a combination of both, 1 only visualized a clock face, and 2 did not know.

Weights and sizes: Finally, for weights and sizes, 21 participants (81%) visualized numbers as digits, 3 (12%) both in digits and letters, and 1 in letters only. One did not know.

3.7. *TTS and the two spelling routes*

When hearing novel words for the first time, we ascribe them a phonological representation, and we can spell them according to the statistics of phoneme–grapheme correspondence. Parallel to this "surface" spelling route, upon hearing a known word, a "lexico-semantic" route allows for the retrieval of the stored orthography, which is particularly required for exception words whose spelling does not obey usual phoneme–grapheme correspondence (e.g. PINT or HAVE in English), and homophone words, whose correct orthography depends on the semantic and syntactic context (e.g. STAIR-STARE).

Novel words: According to the introspective questionnaire, novel words (e.g. brand names or family names) triggered TTS in 20 participants (77%), while 5 (19%) visualized only few letters, and 1 had no synesthesia. Participants were presented with recordings of 10 pseudowords. All participants experienced TTS, which occurred in 248/260 (95%) trials (e.g. most of participants see SOLQUIN for [sɔlk], or see GARMULE for [gaʁmyl]). In 25 participants, the synesthesia was subjectively as automatic and fast as for real words.

Exception words: The 26 participants were presented with recordings of 9 exception words (a tenth one was excluded due to poor recording quality), and they all experienced TTS. We excluded from further analyses 2 participants who did not report the orthographic content of the subtitles. On 167/216 (77%) trials, participants perceived words in their correct spelling. On 32/216 (15%) trials, they perceived an phonologically plausible but lexically incorrect spelling (e.g. YOTE instead of YACHT; RESPE instead of RESPECT). Occasionally, (9/216 trials, 4%) both the correct and an incorrect spelling were perceived (e.g. LONGTEMPS and LONTANT; FEMME and FAME; YOTE then turning into YACHT). There was thus a strong preponderance of lexical spelling, followed by a substantial minority of trials with phonological spelling. A few participants showed an atypical pattern.

Homophones: Finally, participants were presented with 10 homophone words. One participant did not answer. On most trials (189/250, 76%) participants perceived only a single spelling of the target. Some participants sometimes saw more than one spelling, either in succession or simultaneously. In order to formally assess the observation that most responses consisted in relatively frequent nouns, we restricted the analysis to the 189 trials with only one single reported spelling, considering only the 3 most frequently reported spellings for each stimulus. Using a GLME (see Methods) we found that the likelihood for a spelling to be reported was higher if it corresponded to a noun than to another grammatical category ($p = .005$), and to a higher-than to a lower-frequency word ($p = .01$).

Foreign language: Speech perceived in an unfamiliar language may be processed and spelled out in the same way as novel words, provided that a well-defined phonological representation can be computed. According to the introspective questionnaire, when hearing a known foreign language, 17 participants (65%) reported having TTS, 7 (27%) had TTS but to a lesser extent than in their native language, and 2 (8%) had no synesthesia at all. When hearing an unknown language, TTS was reduced to a few sparse letters in 16 participants (62%), occurred roughly like in the usual language in 5 (19%) and did not occur at all in 5 other participants.

3.8. *Synesthesia for non-speech sounds*

Even vocal (e.g. laughing) or non-vocal noises (e.g. explosion) may be spelled out as onomatopoeias. We therefore studied whether such auditory stimuli would also trigger TTS.

According to the introspective questionnaire, at least some non-speech sounds triggered TTS in 14/26 participants (54%), including non-human noises (e.g. bird songs or the sound of a bell), and human noises (e.g. laugh or sneeze). In 11 participants (42%), none of these stimuli triggered TTS. One participant did not know. Participants were presented with 13 recordings of animals, human non-speech sounds, and noises of objects. Animal sounds triggered TTS on 69/104 (66%) trials. All but two participants had a visual perception for at least one sound, ranging from 9/26 for the rooster (seen e.g. as “HUHUHU”), to 24/26 for a cat meowing (generally perceived as the onomatopoeia “MIAOU”). Human noises triggered TTS on 57/104 (55%) trials. All participants experienced TTS for at least one human noise, ranging from 9/26 for a cough (seen e.g. as “HEU—HEU” or “KHU”) to 26/26 for laughs (generally perceived as variants of “AHAHAH”). Noises of objects triggered TTS on only 42/130 (32%) trials. Only 18 participants reported TTS, ranging from 5/26 for a bell sound or a gunshot (one participant saw “TIR”, the French word for shot) to 13/26 for an alarm clock or an ascending musical scale (most of them seeing the words DO RE MI etc.).

4. Discussion

This study is a first attempt at systematically gathering descriptive data on TTS, and at providing an overview of its spectrum. This is necessary ground work, preliminary to subsequent studies of TTS, rather than the test of detailed predictions. Our only background hypothesis is that TTS reflects an enhanced top-down influence of speech processing on orthographic representations, in the form of vivid and automatic mental images of written words. We will now summarize our main findings, relate them to general concepts in the fields of synesthesia and written language processing, and put forward some specific questions and propositions for future studies.

4.1. *Is TTS an exceptional phenomenon?*

Over a few months, we identified 26 synesthetes with reasonable ease. Although we cannot provide a reliable estimate of TTS prevalence in the general population, this suggests that TTS may be in the range of a few percent, similar to other better documented synesthesias. The prevalence of 7% suggested by Chun and Hupé (2013) would place TTS above perfect pitch (.06%, Profita et al., 1988) and grapheme-color synesthesia (1.2%, Carmichael et al., 2015), and below sequence-space (11%, Sagiv et al., 2006) synesthesias. However, the 7% figure is probably a gross overestimation, due to sampling biases. Still, the intuition that TTS would be a highly exceptional phenomenon is clearly misleading, as first suggested by Holm et al. (2015). As most synesthesias, TTS is purely subjective, it is not associated with outstanding abilities or incapacitating impairments, and its existence is generally ignored. Hence it fails to be identified by either synesthetes or their acquaintances. Indeed, in our study, more than 2/3 of participants became aware of the peculiarity of TTS during adulthood, or even only upon reading our recruitment poster.

We found a non-significant 2:1 female/male ratio among participants, which falls in the broad span of previous estimates, ranging from 6:1 (Baron-Cohen et al., 1996; Rich et al., 2005) to 1:1 (Simner et al., 2006). Meta-analyses suggest that sex differences are actually fragile (Simner & Carmichael, 2015) and might be attributed to a self-referral bias, with a higher propensity of females to self-report atypical experiences in general (Dindia & Allen, 1992).

Among our participants, more than 2/3 reported associated synesthesias, space-time synesthesia being the most frequent (52%). Such multiplicity is commonly reported in the synesthesia literature, and is suggestive of common underlying developmental mechanisms (Ward, 2013). Similar to Holm et al. (2015), we found no outstanding link between TTS and grapheme-color synesthesia, which was still present in 6 (23%) of participants. Note that we considered perfect pitch as a type of synesthesia, although this condition is more often considered in the context of music perception (for reviews see Carden & Cline, 2019; Hou et al., 2017). Although this stance may be disputable (for a detailed comparative discussion see Glasser, 2021), perfect pitch fits the definition of synesthesia as stimulation in one sensory or cognitive domain (pitch perception) automatically inducing an unusual and vivid perception in a distinct sensory or cognitive domain (note label). Beyond anecdotal associations in individual persons (e.g. Bernard, 1986; Bouvet et al., 2014; Lebeau et al., 2020), this view is supported by a high frequency of other synesthesias (about 20%) (Gregersen et al., 2013) in individuals with perfect pitch, possibly reflecting phenotypic and genetic relationship (Gregersen et al., 2013; Hänggi et al., 2008). Although the number of participants in the present study is limited, and the prevalence of perfect pitch is highly variable across studies and populations (Glasser, 2021), a proportion of 12% of perfect pitch among our 26 TTS participants, may be viewed with caution as suggestive of an increased incidence.

We also found that about 1/3 of our participants had cases of TTS in their close family. Again, such family predisposition was observed by Galton (1883) and is in line with more recent evidence. Thus Ward and Simner (2005) reported a 16% prevalence inside the families of synesthetes, as compared to 4.4% in the general population (Simner et al., 2006), pointing to the importance of genetic factors (Asher & Carmichael, 2013).

Future controlled studies of prevalence, gender ratio, family history, association with other synesthesias or developmental disorders, and genetics, should allow for a better understanding of the origins of TTS.

4.2. *The paths from sounds to letters in TTS*

TTS is by definition triggered by the perception of speech. What are however the features of speech which are necessary and sufficient, and what are those that modulate the content of TTS subtitles?

Internal phonology. All participants experienced TTS when hearing and watching someone else speaking. In the vast majority of cases, TTS persisted when the speaker remained out of sight, or when participants were speaking themselves. However, what was critical was not the presence of overt speech, but the activation of an internal phonological code. Thus, TTS was usually triggered also by inner speech, a covert phonological stream with no acoustic counterpart (Levelt, 1989). Moreover phonologically irrelevant factors, such as the gender of the speaker, had generally no impact on TTS. The efficiency of only "thinking about" the inducer may actually be a common feature of many synesthesias. Thus most grapheme-color synesthetes assert that letters or numbers that they imagine have the same color as when they see them on the page (Edquist et al., 2006; see also Gould et al., 2014). The abstract nature of the inducer was nicely illustrated by Mroczko et al. (2009): in graphene-color synesthesia, the color of letters is immediately transmitted to novel letter shapes upon learning their graphemic value. We therefore predict that brain imaging should reveal a common activation pattern during TTS, irrespective of the overt or covert type of speech input.

Interestingly still, a few participants reported that their own speech or the speech of hidden speakers did not generate TTS (Fig. 1). In both circumstances, attention may not be focused on the speech stream: speakers who are out of sight are often irrelevant and should be filtered out; and speakers attend to the communicative intent rather than to the sound content of their own utterance. This may result in weaker TTS, or in under-reporting of TTS under those conditions. An alternative account would refer to the tight links that have been proposed between lip-reading and word reading, and the involvement of the left VOT cortex in both abilities (Campbell et al., 1986; Hannagan et al., 2015). One may speculate that in some synesthetes, generating TTS would require combined top-down input to the VWFA, not only from phonology but also from face perception systems.

Non-lexical phonology. Real words have a phonological representation stored in the mental lexicon, but we also manipulate speech sounds in non-lexical contexts, for instance when hearing or reading aloud novel words. In all participants, TTS was triggered by real words, but also by pseudowords, showing again that phonology is critical, irrespective of its lexical or non-lexical source. Even non-speech sounds sometimes elicited TTS, in-as-much as they underwent phonological coding. Thus, one participant told us that a bird used to sing in his garden early in the morning. After a few days he translated the song into an onomatopoeia. From this moment on, the bird's song triggered the synesthetic perception of the written form of the onomatopoeia. A similar logic applies to the triggering of TTS by foreign speech. We therefore predict that the typical brain activation pattern of TTS should appear in the minimal contrast between non-speech noises before vs after learning associated onomatopoeias (Taitz et al., 2018), or between degraded speech before vs after perceptual training (Davis et al., 2005).

From phonology to orthography. The translation from phonology to orthography in TTS followed the parallel pathways featured by the classical two-route model of spelling and reading (Coltheart, 2006; Perry et al., 2007; Taylor et al., 2013). Thus, synesthetic pseudowords were usually spelled according to the most frequent phoneme–grapheme correspondences (phonological route), while exception words were spelled according to their stored orthography. Homophone non-homograph words (e.g. LEAK vs LEEK) were perceived more often under their most frequent orthography, a further trace of lexical processing.

Moreover, when TTS was triggered by non-speech noises, TTS occasionally disclosed a contribution of semantic knowledge. First, the content of the TTS sometimes corresponded to a largely conventional phonological code. Thus the "DRIIIING" spelling often elicited by the sound of an alarm clock may not be its most faithful translation, but it is its usual literary depiction, e.g. in comic strips. Meowing elicited the very familiar "MIAOU" in almost all participants, while the less familiar call of a rooster was a less effective inducer. The same trend prevailed for other categories of noise: laughter was more effective than cough, and alarm-clock ringing than gunshot. Other types of synesthesias also are more vivid for more frequent inducer–concurrent associations (Beeli et al., 2007). Second, some participants, instead of seeing an onomatopoeia, visualized the written name of the object generating the noise. For example, some participants saw the French word for "GUN" or "GUNSHOT" when hearing the noise of a gunshot. It is likely that the synesthete identified the noise, internally produced the corresponding word, which then entered in the common TTS pathway.

The case of number words. In addition to generic orthographic coding (THIRTY), numbers can be written using Arabic numerals (30). Across participants, there was a remarkable variety in the relative contribution of the two codes to TTS, and mixtures of both were not uncommon (e.g. one participant reported when asked about the date that he saw "the numbers of the day in Arabic numerals, months in letters, and year in numerals"; another reported about the number 35 that he saw "letters and numerals overlapping"). This may be similar to the collisions occurring with exception words, which sometimes induced the near simultaneous perception of the lexical and phonological spellings (e.g. FEMME and FAME). Such collisions indicate that TTS subtitles can be fed with the output of whatever route linking phonology to orthography.

In summary, TTS resulted from the translation from phonology to orthography through phonological and lexico-semantic routes. Importantly, even neurotypical individuals show behavioral evidence of phonological influence on visual word recognition. First, there is implicit and automatic activation of orthographic codes during speech perception. For example, during purely auditory lexical decision, latencies are shorter for words with a consistent than with an inconsistent spelling (Ziegler et al., 2008), an effect which disappears under TMS of the SMG (Pattamadilok et al., 2010). Similarly, also during auditory lexical decision, prime words facilitate targets in proportion to their orthographic overlap (Chéreau et al., 2007). Second, phonology does actually modulate the recognition of printed words. For instance, during lexical decision and naming with visual words, responses are facilitated by primes with a purely phonological link to targets (Carreiras et al., 2005; Grainger et al., 2006). In this context, TTS may be viewed as an exceptionally vivid access to orthographic representations which are activated even in non-synesthetes, implying some continuity between normal speech perception and TTS, a controversial issue in the field of synesthesia studies (Deroy & Spence, 2016; Martino & Marks, 2001).

In a recent single-case fMRI study of TTS, we used dynamic causal modeling and showed that indeed TTS induced by spoken words and pseudowords relied on the top-down flow of information along distinct routes, involving the supramarginal and middle temporal gyri, respectively (Hauw et al., 2022). Future studies should resort to time resolved methods such as EEG/MEG in order to decipher the temporal course of the alternative processes leading to TTS.

4.3. *Visual features of TTS*

All participants were able to report with some precision the core visual features of TTS in terms of location, distance, color, or case. While orthographic imagery exists to some degree in all literate individuals (Bourlon et al., 2009; Holm et al., 2015), tickertapers feature a particularly vivid form of such imagery, possibly correlated with higher activations in the VWFA (Fulford et al., 2018; Hauw et al., 2022).

The precise features of synesthesias (e.g. what is a the color of a specific letter?) are at least partially determined by personal experience (Rich et al., 2005; Ward & Simner, 2003; Witthoft & Winawer, 2006). In the case of TTS, the black color of subtitles may reflect the usual color of printed words, and their location at the bottom of the visual field or next to the speaker's mouth may result from the position of actual subtitles and of comic bubbles.

Four participants perceived colored TTS subtitles. They all had associated TTS and grapheme-color synesthesia. To their mind's eye, letters in the subtitles were colored in the same way as external letters, revealing the chaining of the two synesthesias, mapping phonology to orthography, and then orthography to colors. In 9 other participants, the visual features of subtitles were modulated by the emotional content of the input. In various synesthesias also, emotions elicited by the inducer may modulate the features of the concurrent (for a review, see Dael et al., 2013). It may even occur that “emotional states as such exist as colors projected out in the visual field, (…) and the intensity and prevalence the color are a measure of the intensity of the emotional reaction” (Cutsforth, 1925). This may suggest that in our 4 participants, TTS was associated with a putative “affective synesthesia”.

In grapheme-color synesthesia, distinction is commonly made between a majority of “associators” and a minority of “projectors”, who feel that illusory colors exist in their mind or in the outside world, respectively (Amsel et al., 2017; Dixon et al., 2004). Although the distinction may not be always clear-cut, we found a similar predominance of TTS associators (69%). This distinction putatively results from differences in the anatomical connectivity among involved brain areas (Rouw & Scholte, 2007), or in the pattern of their functional cross-talk (Amsel et al., 2017; van Leeuwen et al., 2011). Assuming that TTS associators and projectors put into play regions involved in mental imagery and perception, respectively (Pearson, 2019; Spagna et al., 2021), one may predict that the two groups should differ in the relative involvement of high-level infero-temporal areas such as the VWFA, vs. of low-level visual cortex.

In future research it may also be interesting to investigate the visuo-spatial arraying of the alphabet in the mental imagery of TTS individuals (Jonas et al., 2011), as it may potentially differ from participants with typical relationships between phonology and orthography.

4.4. *Dynamic features of TTS*

Qualitative introspective reports are obviously inadequate to characterize the fast time course of TTS. One may only observe that the life span of images is at most in the order of seconds, that they are rapidly flushed by the following ones, and that they encompass from one to a few words. Those features are suggestive of a short term memory system, possibly identical to the so-called graphemic buffer, a device which has been studied mostly in the context of word spelling. It is thought to maintain an orthographic representation of the current word during writing or typing (Buchwald & Rapp, 2004), but it is also involved to some extent in reading (Caramazza et al., 1996; Purcell et al., 2017). In this context, TTS might be considered as the result of a hyperactive graphemic buffer, supporting enhanced visual images. This could indeed explain why 10 (38%) participants considered that being endowed with TTS facilitated writing. Future studies should study the links between TTS and writing performance, and resort to time-resolved methods to assess the time course of the processes underlying TTS.

4.5. *The pros and cons of TTS*

Easier writing and better orthography is the only advantage consistently acknowledged by tickertapers. In grapheme-color synesthesia, which like TTS is closely dependent on literacy, synesthetes seem to benefit from a modest but replicable advantage in verbal and other memory tasks (for a review see Rothen et al., 2012). More recently, Smees et al. (2019) showed that 10–11 year-old synesthetes have better expressive and receptive vocabulary than control children. This advantage did not exist before elementary school, suggesting that synesthesia yielded a verbal learning advantage. In the case of TTS, one may predict that effortlessly enjoying a stable and vivid orthographic representation should help tickertapers to learn to read and write.

Conversely, most participants reported that TTS makes it difficult for them to focus their attention, for example when trying to read in public spaces, or to follow one among several ongoing conversations. Attentional effort allowed some of them to get rid of the irrelevant channel, for instance by putting actual printed words and concurrent TTS subtitles in different “mental spaces”.

In summary, factual evidence is missing on the practical pros and cons of TTS. We predict that tickertapers should be better than controls in tasks based on stable orthographic imagery, such as writing long and complex words, or letter counting, backward spelling, or visual rhyme judgment with spoken words. Conversely, concurrent speech should interfere with e.g. visual lexical decision more detrimentally in tickertapers than in controls.

5. Conclusion

Generally considered an anecdotal and exceptional phenomenon, tickertape synesthesia provides a novel gate in the mechanisms of reading and reading acquisition. Similar to dyslexia, it results from an atypical setup of the cerebral reading machinery. We reported a descriptive study of tickertape synesthesia in a group of 26 participants, based on detailed introspective questionnaires and on the presentation of controlled auditory stimuli. This groundwork leaves a number of questions open for future research. Empirical studies should address (1) the genetic and environmental bases of TTS; (2) the objective behavioral markers of TTS. This is particularly important, not only for the understanding of the cognitive underpinnings of TTS, but also for developing objective criteria for ensuring the genuineness of TTS beyond subjective self-report; (3) the development of orthographic imagery and the emergence of TTS in children with or without dyslexia; (4) the cerebral mechanisms of TTS using multimodal brain imaging in adults and children; (5) the brain correlates of individual variability in TTS phenotype.

Credit author statement

Hauw Fabien: Investigation, Writing – Original Draft, Writing - Review & Editing.
El Soudany Mohamed: Conceptualization, Methodology, Data curation.
Cohen Laurent: Conceptualization, Methodology, Writing – Original Draft, Writing- Reviewing and Editing, Supervision.

Funding statement

This study was funded by the “Investissements d'avenir” program (ANR-10-IAHU-06) to the Paris Brain Institute, the “TOPLEX” ANR program to Laurent Cohen, and the “Fondation pour la Recherche Médicale” to Fabien Hauw.

Open practices

The study in this article earned Open Data and Open Materials badges for transparent practices. Materials and data for the study are available at https://zenodo.org/record/7108817#

Declaration of competing interest

FH, MES, & LC disclose no competing interests.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2022.11.005.

REFERENCES

Amsel, B. D., Kutas, M., & Coulson, S. (2017). Projectors, associators, visual imagery, and the time course of visual processing in grapheme-color synesthesia. *Cognitive Neuroscience, 8*, 206–223. https://doi.org/10.1080/17588928.2017.1353492

Asher, J. E., & Carmichael, D. A. (2013). The genetics and inheritance of synesthesia. *Oxford Handbook of Synesthesia*, 23–45.

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia: Prevalence and familiality. *Perception, 25*, 1073–1079. https://doi.org/10.1068/p251073

Beeli, G., Esslen, M., & Jäncke, L. (2007). Frequency correlates in grapheme-color synaesthesia. *Psychological Science, 18*, 788–792. https://doi.org/10.1111/j.1467-9280.2007.01980.x

Bernard, J. W. (1986). Messiaen's synaesthesia: The correspondence between color and sound structure in his music. *Music Perception, 4*, 41–68.

Bourlon, C., Chokron, S., Bachoud-Lévi, A.-C., Coubard, O., Bergeras, I., Moulignier, A., Viret, A. C., & Bartolomeo, P. (2009). Presentation of an assessment battery for visual mental imagery and visual perception. *Revue Neurologique, 165*, 1045–1054.

Bouvet, L., Donnadieu, S., Valdois, S., Caron, C., Dawson, M., & Mottron, L. (2014). Veridical mapping in savant abilities, absolute pitch, and synesthesia: An autism case study. *Frontiers in Psychology, 5*, 106.

Buchwald, A., & Rapp, B. (2004). Rethinking the graphemic buffer? *Brain and Language, 91*, 100–101.

Campbell, R., Landis, T., & Regard, M. (1986). Face recognition and lipreading. A neurological dissociation. *Brain: a Journal of Neurology, 109*(Pt 3), 509–521. https://doi.org/10.1093/brain/109.3.509

Caramazza, A., Capasso, R., & Miceli, G. (1996). The role of the graphemic buffer in reading. *Cognitive Neuropsychology, 13*, 673–698.

Carden, J., & Cline, T. (2019). Absolute pitch: Myths, evidence and relevance to music education and performance. *Psychology of Music, 47*, 890–901.

Carmichael, D. A., Down, M. P., Shillcock, R. C., Eagleman, D. M., & Simner, J. (2015). Validating a standardised test battery for synesthesia: Does the synesthesia battery reliably detect synesthesia? Conscious. *Cogn, 33*, 375–385. https://doi.org/10.1016/j.concog.2015.02.001

Carreiras, M., Ferrand, L., Grainger, J., & Perea, M. (2005). Sequential effects of phonological priming in visual word recognition. *Psychological Science, 16*, 585–589. https://doi.org/10.1111/j.1467-9280.2005.01579.x

Chéreau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming. *Cognition, 102*, 341–360. https://doi.org/10.1016/j.cognition.2006.01.001

Chun, C. A., & Hupé, J.-M. (2013). Mirror-touch and ticker tape experiences in synesthesia. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00776

Coltheart, M. (2006). Dual route and connectionist models of reading: An overview. *London Review of Education*. https://doi.org/10.1080/13603110600574322

Coltheart, M., & Glick, M. J. (1974). Visual imagery: A case study. *Quarterly Journal of Experimental Psychology, 26*, 438–453. https://doi.org/10.1080/14640747408400433

Cutsforth, T. D. (1925). The rôle of emotion in a synaesthetic subject. *The American Journal of Psychology, 36*, 527–543. https://doi.org/10.2307/1413908

Dael, N., Sierro, G., & Mohr, C. (2013). Affect-related synesthesias: A prospective view on their existence, expression and underlying mechanisms. *Frontiers in Psychology, 4*.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of*

Experimental Psychology. General, 134, 222–241. https://doi.org/10.1037/0096-3445.134.2.222
Dehaene-Lambertz, G., Monzalvo, K., & Dehaene, S. (2018). The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition. *Plos Biology, 16*, Article e2004103. https://doi.org/10.1371/journal.pbio.2004103
Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews. Neuroscience, 16*, 234–244. https://doi.org/10.1038/nrn3924
Deroy, O., & Spence, C. (2016). Lessons of synaesthesia for consciousness: Learning from the exception, rather than the general. *Neuropsychologia, 88*, 49–57. https://doi.org/10.1016/j.neuropsychologia.2015.08.005
Dindia, K., & Allen, M. (1992). Sex differences in self-disclosure: A meta-analysis. *Psychological Bulletin, 112*, 106–124. https://doi.org/10.1037/0033-2909.112.1.106
Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal: Projector versus associator synaesthetes. *Cognitive, Affective & Behavioral Neuroscience, 4*, 335–343. https://doi.org/10.3758/CABN.4.3.335
Edquist, J., Rich, A., Brinkman, C., & Mattingley, J. (2006). Do synaesthetic colours act as unique features in visual search? *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 42*, 222–231. https://doi.org/10.1016/S0010-9452(08)70347-2
Friedmann, N., & Coltheart, M. (2016). Types of developmental dyslexia. Handbook of communication disorders. *Theoretical, empirical, and applied linguistics perspectives*, 1–37.
Fulford, J., Milton, F., Salas, D., Smith, A., Simler, A., Winlove, C., & Zeman, A. (2018). The neural correlates of visual imagery vividness – An fMRI study and literature review. *Cortex, The Eye's Mind - visual imagination, neuroscience and the humanities, 105*, 26–40. https://doi.org/10.1016/j.cortex.2017.09.014
Galton, F. (1883). *Inquiries into human faculty and its development.* Macmillan.
Glasser, S. (2021). Perceiving music through the lens of synaesthesia and absolute pitch. *Perception, 50*, 690–708.
Gould, C., Froese, T., Barrett, A. B., Ward, J., & Seth, A. K. (2014). An extended case study on the phenomenology of sequence-space synesthesia. *Frontiers in Human Neuroscience, 8*. https://doi.org/10.3389/fnhum.2014.00433
Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The time course of orthographic and phonological code activation. *Psychological Science, 17*, 1021–1026.
Gregersen, P. K., Kowalsky, E., Lee, A., Baron-Cohen, S., Fisher, S. E., Asher, J. E., Ballard, D., Freudenberg, J., & Li, W. (2013). Absolute pitch exhibits phenotypic and genetic overlap with synesthesia. *Human Molecular Genetics, 22*, 2097–2104. https://doi.org/10.1093/hmg/ddt059
Hänggi, J., Beeli, G., Oechslin, M. S., & Jäncke, L. (2008). The multiple synaesthete ES—Neuroanatomical basis of interval-taste and tone-colour synaesthesia. *NeuroImage, 43*, 192–203.
Hannagan, T., Amedi, A., Cohen, L., Dehaene-Lambertz, G., & Dehaene, S. (2015). Origins of the specialization for letters and numbers in ventral occipitotemporal cortex. *Trends in Cognitive Sciences, 19*, 374–382. https://doi.org/10.1016/j.tics.2015.05.006
Harlaar, N., Dale, P. S., & Plomin, R. (2007). From learning to read to reading to learn: Substantial and stable genetic influence. *Child Development, 78*, 116–131. https://doi.org/10.1111/j.1467-8624.2007.00988.x
Hauw, F., El Soudany, M., Rosso, C., Daunizeau, J., & Cohen, L. (2022). Seeing speech: The cerebral substrate of tickertape synesthesia. (preprint). *Neuroscience*. https://doi.org/10.1101/2022.09.19.508477
Holm, S., Eilertsen, T., & Price, M. C. (2015). How uncommon is tickertaping? Prevalence and characteristics of seeing the words you hear. *Cognitive Neuroscience, 6*, 89–99. https://doi.org/10.1080/17588928.2015.1048209
Hou, J., Chen, A. C., Song, B., Sun, C., & Beauchaine, T. P. (2017). Neural correlates of absolute pitch: A review. *Music & Scienc, 21*, 287–302.
Johnson, D., Allison, C., & Baron-Cohen, S. (2013). The prevalence of synesthesia. *Oxford Handbook of Synesthesia, 1*.
Jonas, C. N., Taylor, A. J. G., Hutton, S., Weiss, P. H., & Ward, J. (2011). Visuo-spatial representations of the alphabet in synaesthetes and non-synaesthetes: Visuo-spatial alphabets. *Journal of Neuropsychology, 5*, 302–322. https://doi.org/10.1111/j.1748-6653.2011.02010.x
Lebeau, C., Tremblay, M. N., & Richer, F. (2020). Adaptation to a new tuning standard in a musician with tone-color synesthesia and absolute pitch. *Auditory Perception & Cognition, 3*, 113–123.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge MA: MIT Press.
Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science, 10*, 61–65.
Mascheretti, S., Andreola, C., Scaini, S., & Sulpizio, S. (2018). Beyond genes: A systematic review of environmental risk factors in specific reading disorder. *Research in Developmental Disabilities, 82*, 147–152. https://doi.org/10.1016/j.ridd.2018.03.005
Mroczko, A., Metzinger, T., Singer, W., & Nikolic, D. (2009). Immediate transfer of synesthesia to a novel inducer. *Journal of Visualization, 9*. https://doi.org/10.1167/9.12.25, 25–25.
Norton, E. S., Beach, S. D., & Gabrieli, J. D. E. (2015). Neurobiology of dyslexia. *Current Opinion in Neurobiology, 30*, 73–78. https://doi.org/10.1016/j.conb.2014.09.007
Pattamadilok, C., Knierim, I. N., Kawabata Duncan, K. J., & Devlin, J. T. (2010). How does learning to read affect speech perception? *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 30*, 8435–8444. https://doi.org/10.1523/JNEUROSCI.5791-09.2010
Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews. Neuroscience, 20*, 624–634. https://doi.org/10.1038/s41583-019-0202-9
Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114*, 273–315. https://doi.org/10.1037/0033-295X.114.2.273
Profita, J., Bidder, T. G., Optiz, J. M., & Reynolds, J. F. (1988). Perfect pitch. *American Journal of Medical Genetics, 29*, 763–771. https://doi.org/10.1002/ajmg.1320290405
Purcell, J. J., Jiang, X., & Eden, G. F. (2017). Shared orthographic neuronal representations for spelling and reading. *NeuroImage, 147*, 554–567. https://doi.org/10.1016/j.neuroimage.2016.12.054
Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., & Scotto di Covella, L. (2018). Neuroanatomy of developmental dyslexia: Pitfalls and promise. *Neuroscience and Biobehavioral Reviews, 84*, 434–452. https://doi.org/10.1016/j.neubiorev.2017.08.001
Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia: Implications for the role of early experience in lexical-colour associations. *Cognition, 98*, 53–84. https://doi.org/10.1016/j.cognition.2004.11.003
Rothen, N., & Meier, B. (2010). Higher prevalence of synaesthesia in art students. *Perception, 39*, 718–720. https://doi.org/10.1068/p6680
Rothen, N., Meier, B., & Ward, J. (2012). Enhanced memory ability: Insights from synaesthesia. *Neuroscience and Biobehavioral Reviews, 36*, 1952–1963. https://doi.org/10.1016/j.neubiorev.2012.05.004
Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience, 10*, 792–797. https://doi.org/10.1038/nn1906
Sagiv, N., Simner, J., Collins, J., Butterworth, B., & Ward, J. (2006). What is the relationship between synaesthesia and visuo-

spatial number forms? *Cognition, 101*, 114–128. https://doi.org/10.1016/j.cognition.2005.09.004

Simner, J., & Carmichael, D. A. (2015). Is synaesthesia a dominantly female trait? *Cognitive Neuroscience, 6*, 68–76. https://doi.org/10.1080/17588928.2015.1019441

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., Scott, K., & Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception, 35*, 1024–1033. https://doi.org/10.1068/p5469

Smees, R., Hughes, J., Carmichael, D. A., & Simner, J. (2019). Learning in colour: Children with grapheme-colour synaesthesia show cognitive benefits in vocabulary and self-evaluated reading. *Philosophical Transactions of the Royal Society B: Biological Sciences, 374*, Article 20180348. https://doi.org/10.1098/rstb.2018.0348

Spagna, A., Hajhajate, D., Liu, J., & Bartolomeo, P. (2021). Visual mental imagery engages the left fusiform gyrus, but not the early visual cortex: A meta-analysis of neuroimaging evidence. *Neuroscience and Biobehavioral Reviews, 122*, 201–217. https://doi.org/10.1016/j.neubiorev.2020.12.029

Taitz, A., Assaneo, M. F., Elisei, N., Trípodi, M., Cohen, L., Sitt, J. D., & Trevisan, M. A. (2018). The audiovisual structure of onomatopoeias: An intrusion of real-world physics in lexical creation. *PLoS One, 13*, Article e0193466.

Taylor, J. S., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin Journal, 139*(4), 766–791.

van Leeuwen, T. M., den Ouden, H. E. M., & Hagoort, P. (2011). Effective connectivity determines the nature of subjective experience in gGrapheme-color synesthesia. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 31*, 9879–9884. https://doi.org/10.1523/JNEUROSCI.0569-11.2011

Ward, J. (2013). Synesthesia. *Annual Review of Psychology, 64*, 49–75. https://doi.org/10.1146/annurev-psych-113011-143840

Ward, J., & Simner, J. (2003). Lexical-gustatory synaesthesia: Linguistic and conceptual factors. *Cognition, 89*, 237–261. https://doi.org/10.1016/S0010-0277(03)00122-7

Ward, J., & Simner, J. (2005). Is synaesthesia an X-linked dominant trait with lethality in males? *Perception, 34*, 611–623. https://doi.org/10.1068/p5250

Witthoft, N., & Winawer, J. (2006). Synesthetic colors determined by having colored refrigerator magnets in childhood. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 42*, 175–183. https://doi.org/10.1016/S0010-9452(08)70342-3

Ziegler, J. C., Université, A., Recherche, C. N. D. L., Petrova, A., Ferrand, L., Pascal, U. B., National, C., & Scientifique, R. (2008). Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 643–661.