

Multisensory cueing facilitates naming in aphasia

Klaudia Grechuta (✉ klaudiagrechuta@gmail.com)

Institut de Bioenginyeria de Catalunya <https://orcid.org/0000-0002-4954-1372>

Belén Rubio Ballester

Institut de Bioenginyeria de Catalunya

Rosa Espín Munné

Hospital Universitari de Tarragona Joan XXIII

Teresa Usabiaga Bernal

Hospital Universitari de Tarragona Joan XXIII

Begoña Molina Hervás

Hospital Universitari de Tarragona Joan XXIII

Bettina Mohr

Charite Universitätsmedizin Berlin

Friedemann Pulvermüller

Freie Universität Berlin

Rosa San Segundo

Hospital Universitari de Tarragona Joan XXIII

Paul FMJ Verschure

Institut de Bioenginyeria de Catalunya

Research

Keywords: Stroke, Aphasia, Lexical access, Word-finding, Multisensory cueing, Neurorehabilitation

Posted Date: August 15th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-28744/v2>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on September 9th, 2020. See the published version at <https://doi.org/10.1186/s12984-020-00751-w>.

Abstract

Background: Impaired naming is a ubiquitous symptom in all types of aphasia, which often adversely impacts independence, quality of life, and recovery of affected individuals. Previous research has demonstrated that naming can be facilitated by phonological and semantic cueing strategies that are largely incorporated into the treatment of anomic disturbances. Beneficial effects of cueing, whereby naming becomes faster and more accurate, are often attributed to the priming mechanisms occurring within the distributed language network.

Objective: We proposed and explored two novel cueing techniques: (1) Silent Visuomotor Cues (SVC), which provided articulatory information of target words presented in the form of silent videos, and (2) Semantic Auditory Cues (SAC), which consisted of acoustic information semantically relevant to target words (ringing for “telephone”). Grounded in neurophysiological evidence, we hypothesized that both SVC and SAC might aid communicative effectiveness possibly by triggering activity in perceptual and semantic language regions, respectively.

Methods: Ten participants with chronic non-fluent aphasia were recruited for a longitudinal clinical intervention. Participants were split into dyads (i.e., five pairs of two participants) and required to engage in a turn-based peer-to-peer language game using the Rehabilitation Gaming System for aphasia (RGSa). The objective of the RGSa sessions was to practice communicative acts, such as making a request. We administered SVCs and SACs in a pseudorandomized manner at the moment when the active player selected the object to be requested from the interlocutor. For the analysis, we compared the times from selection to the reception of the desired object between cued and non-cued trials.

Results: Naming accuracy, as measured by a standard clinical scale, significantly improved for all stimuli at each evaluation point, including the follow-up. Moreover, the results yielded beneficial effects of both SVC and SAC cues on word naming, especially at the early intervention sessions when the exposure to the target lexicon was infrequent.

Conclusions: This study supports the efficacy of the proposed cueing strategies which could be integrated into the clinic or mobile technology to aid naming even at the chronic stages of aphasia. These findings are consistent with sensorimotor accounts of language processing, suggesting a coupling between language, motor, and semantic brain regions.

Trial registration: NCT02928822. Registered 30 May 2016

1. Introduction

About 30% of stroke patients worldwide suffer from language disorders, and the majority remains chronic [1]. Anomia, or word-finding difficulty, is a ubiquitous characteristic of aphasia, which significantly

compromises communication and quality of life of individuals affected by stroke [2], [3]. Consequently, aphasia rehabilitation largely incorporates strategies fostering the recovery of impaired naming and communication by facilitating access to linguistic content.

To map a lexical concept to verbal structure requires multiple steps [4], [5]. First, there is the intention to articulate a specific concept in speech, followed by the so-called lexical access, which consists of the retrieval of a target word from a lexicon [6], [7]. At this stage, the focused concept activates the target lemma, the semantic and syntactic properties of the lexical item [8], triggering the speech form-defining, phonological system. The latter provides verbal execution where the articulatory shape of a word, in the context of other words, forms a sentence-like utterance [9]. In stroke-induced aphasia, depending on the lesion site and extent, some or all stages of this naming process might be impaired, leading to high variability in language performance deficits among affected individuals. Consequently, standard naming therapy, or the so-called cueing, is designed to address different phases and aspects of both retrieval and production [3]. For example, the well-established phonological cueing approach targets the ability to retrieve phonemes underlying the articulation of a word [10], [11]. To this aim, patients are given verbal cues that provide initial sound/s of the target word (e.g., “p” for “pancake”). Another therapeutic method is the semantic cueing, which targets the activation of lexical-semantic association networks [12], [13]. As such, semantic cueing consists of providing information that categorizes, describes, or defines target words (e.g., “it goes well with maple syrup” for a pancake).

In the clinical context, cueing is considered beneficial because it facilitates naming, consequently resulting in higher accuracy and faster reaction times of speech production. Indeed, phonological, semantic, and mixed approaches substantially improve not only immediate but also long-term naming performance as well as functional communicative effectiveness [14]–[18]. Critically, similar effects are reported when the cues are administered through technology-based methods, even to individuals with persisting aphasia [19]–[23]. This finding is particularly relevant in the context of rapid advancement of self-managed, computer-based exercises for individuals with aphasia, which are becoming widely tested and used not only as a part of the clinical inpatient care during the acute and subacute stages but also after the hospital discharge at patient’s homes [24].

The beneficial effects of cueing, whereby the naming of target words becomes faster and more accurate, are usually attributed to priming mechanisms occurring within residual language network bilaterally [25]–[27]. Depending on the type of administered cues (e.g., initial phoneme, full word), imaging studies report increased activity in regions including the right anterior insula, inferior frontal, and dorsal anterior cingulate cortices, as well as the left premotor cortex [28]. One account yields that cueing elicits activation of lexical representations at phonological and semantic levels in a selective manner [29], thus enabling the recovery of phonological or semantic deficits, respectively. This hypothesis, however, seems at odds with the notion that during therapeutic tasks such as picture naming semantic information contained by the stimuli might automatically activate phonological information and vice-versa [30], [31]. This interpretation might be explained by the interactive activation approach to word production, which proposes that lexical retrieval occurs within a distributed language network, in which nodes are connected

across semantic, lexical, and phonological levels of representation in a feedforward and feedback manner (i.e., bidirectionally) [5]. Indeed, an analysis of the language connectome in both healthy controls and brain tumor patients showed a broad network spanning about 25% of the total human connectome [32]. Following this architecture, therapy-induced stimulation at the level of the semantic system can activate phonological and orthographical processing, and vice-versa. This, in turn, may explain why several studies report higher efficacy of a combined (i.e., mixed) cueing therapy rather than when semantic or phonological primes are delivered independently [18], [30]. Further supporting evidence for this network interpretation is the observation that speech perception is governed by general principles of statistical inference across all available perceptual sources [33] suggesting that similar principles of Bayesian inference are involved in cueing based rehabilitation strategies.

In this study, we aim to test the inference-based network perspective on language and its deficits. To this end, we propose two novel cueing strategies and investigate their effects on naming in the context of a within-subjects longitudinal clinical study with post-stroke aphasia patients. On the one hand, we investigated the so-called Silent Visuomotor Cues (SVC) strategy. SVCs provided articulatory information of target words presented in the form of silent videos which display lip movements of a speech and language therapist during naming [34]. On the other hand, we studied Semantic Auditory Cues (SAC). Here, the primes consisted of acoustic information semantically relevant to target words such as the sound of ringing for “telephone,” or the sound of an engine revving up for “car.”

First, the motivation to investigate SVC was grounded in neurophysiological evidence, which strongly supports the notion of perceptual functions of speech production centers. In particular, it has been demonstrated that part of the ventrolateral frontal cortex in humans (Brodmann’s 44), initially thought to be engaged in the control of motor aspects of speech production exclusively [35], [36], is also involved in the processing of orofacial gestures [37], [38]. This is well illustrated in a MEG (magnetoencephalography) study in which the authors compared the activation of the human Mirror-Neuron System, (MNS) including Broca’s area, during execution, observation and imitation of verbal and nonverbal lip forms [38]. The stimuli were presented in the form of static images illustrating orofacial gestures that solely imply action (i.e., motionless). Interestingly, the results yielded strong BOLD signals evoked bilaterally in the MNS, including Brodmann’s areas 44/45 (Broca’s area), during pure perception of lip forms. This finding explicitly demonstrates that viewing visual orofacial stimuli is sufficient to trigger activity in the distributed language network, including areas involved in word-finding and speech production. We, therefore, hypothesized that providing CVS, that is muted videos presenting lips articulating target word, might improve verbal performance, suggesting improved retrieval in participants with aphasia.

Second, we aimed to empirically explore the effects of SAC on lexical access and verbal execution in the same group. We chose to study whether semantically relevant sounds positively impact naming based on the notion of an embodied inference driven language network, which proposes that auditory and conceptual brain systems are neuroanatomically and functionally coupled [39], [40], driven by the statistics of real-world interaction [41], [42]. Specifically, a functional Magnetic Resonance Imaging (fMRI)

study [39] revealed that cortical activations induced by listening to sounds of objects and animals (e.g., “ringing” or “barking”) overlap with activations induced by merely reading words that contain auditory features (e.g., “telephone”, “dog”). The authors reported the overlap in the posterior superior temporal gyrus (pSTG) and middle temporal gyrus (MTG), which suggests that common neural sources underlie auditory perception and processing of words that comprise acoustic features. Critically, MTG plays a significant role within the brain’s language network during syntactic processing in both comprehension and production of speech [43]. For example, MTG was shown to subserve the retrieval, including selection and integration, of lexical–syntactic information in a syntactic ambiguity resolution task [44], [45]. Interestingly, pSTG is also involved in speech production, which is evidenced by clinical studies of conduction aphasia [46] as well as behavioral and imaging experiments with healthy subjects who performed tasks that included word generation [47], reading [48], syllable rehearsal [49], and naming [50]–[52]. Hence, here we hypothesized that providing aphasia patients with SAC may facilitate naming possibly by activating brain regions involved in language production processing.

Similar to phonological and semantic cueing, we reasoned that, if the proposed SVC and SAC strategies are beneficial for the recovery of anomic disturbances in aphasia, they will foster naming accuracy and communication skills. We delivered and tested the efficacy of both types of cues in the context of longitudinal clinical intervention in which participants underwent a peer-to-peer Virtual Reality (VR)-based language therapy using the Rehabilitation Gaming System for aphasia (RGSa) [22], which incorporates principles of Intensive Language Action Therapy (ILAT) [53], [54].

2. Methods

2.1. Participants

Ten participants with chronic (> 6 months post-stroke, mean (SD): 69.9(48.7)) aphasia participated in the study (age: Mean (SD): 57.6(9.9)). We included participants with moderate-to-severe stages of non-fluent aphasia as identified by a standard screening tool [55]. All participants were right-handed as assessed by the Edinburgh Handedness Inventory [56] and suffered a single left-hemispheric stroke affecting frontotemporal and parietal cortical areas, as evidenced by CT or MRI-scans. Participants were excluded (1) if they had a speech and language disorder caused by a neurological deficit other than stroke, (2) if they had severe and untreated forms of cognitive disorders (assessed by the Mini-Mental State Examination [57]) and motor impairments (determined using Fugl-Meyer Assessment Upper Extremity [58]), which could adversely affect participation in the study and interaction with the proposed system, (3) if two years before the enrollment they participated in alternative intensive interventions, or (4) if they were currently using another computer program that trains naming or word finding. The demographic sample characteristics of all participants are presented in Table 2.

The reported paradigm deployed a within-subjects design. The experimental procedures followed written consents from all the involved participants. The study was further approved by the local Ethical

Committee from the Hospital Universitari Joan XXIII and registered on ClinicalTrials.gov (identifier: NCT02928822) [22]. Clinical results of the randomized controlled trial are reported in [22].

2.2. Treatment protocol and setting

All participants received five weekly intervention-sessions for two months. The duration of each session was 30–40 min. Thus, the full treatment included a total of approximately 23 hours per participant.

The proposed cueing strategies were integrated into a novel language rehabilitation paradigm, the so-called Rehabilitation Gaming System for aphasia (RGSa) [22], [59]. Inspired by Intensive Language Action Therapy (ILAT) [53], RGSa is a VR-based rehabilitation tool administered in the form of an interactive language game that aims at practicing both speech production and comprehension by training frequent and behaviorally relevant communication acts such as making a request [22], [53], [60]. To this end, during therapeutic sessions of RGSa, ten participants were split into dyads (i.e., five pairs of two participants) and required to engage in a turn-based game played in a peer-to-peer setting without the involvement of a therapist [61].

The therapeutic setup of RGSa included two personal computers (Vaio, Japan) connected through a local area network, two headsets (EX-01 BluetoothR, Gioteck, Canada), and two motion tracking sensors (Kinect2, Microsoft, USA). Participants sat in a hospital ward in front of each other facing their respective screens which displayed the virtual environment from the first-person perspective. The virtual scene aimed to represent the actual setting. Thus, it consisted of two avatars seated at the respective sides of the table such that participants could see their virtual arms and the avatar of their interlocutor. On the virtual table, there was a set of three identical objects (see Stimuli) simultaneously available for selection. The movements of the real arms were continuously tracked by the Kinect and mapped in the real-time onto the arms of the virtual avatar. This method enabled interaction with the virtual world and, in particular, the virtual objects.

The objective of each session, and each participant was to collect as many objects as possible by requesting them from the other player or handing them over when required [22], [53]. At the beginning of each (daily) session, one of the participants from a dyad (e.g., PlayerA) was randomly assigned to initiate the game. Every trial consisted of the following three steps:

(1)

PlayerA chooses the desired object. PlayerA indicates the choice of the object for request by reaching towards it. To select the object, players were required to place the avatar's hand over that object for three consecutive seconds. Once selected, to increase saliency and facilitate interaction, the object would light up in yellow and start rotating slowly over the vertical axis. Critically, to test our prediction about the beneficial effects of multisensory cueing on word production, we provided either SVC or SAC immediately after object selection to half of the stimuli (see Stimuli).

(2)

PlayerA verbally requests the matching object from PlayerB. After object selection, PlayerA had to utter a verbal request to obtain the matching object from the opponent. The use of politeness forms and full phrases ("Please, could you pass me the *pancake*") was encouraged but not necessary.

(3)

PlayerB reacts to the request. In case the request was not understood, PlayerB had to ask PlayerA to repeat or clarify the request until it was clear. Whenever PlayerB understood what object was being requested, they were required to hand over the matching object to PlayerA by reaching towards it and holding the virtual hand over it for three consecutive seconds.

The completion of these three steps comprised a successful communicative interaction, which included both a successful speech act (performed by PlayerA) and successful comprehension (performed by PlayerB). After such a sequence of events, both participants saw the two matching objects on the screen (i.e., positive feedback), heard the correct pronunciation of the target word through headphones (i.e., reinforcement), received a point, and the turn changed (Fig. 1A). After a short delay, a new pseudo-randomly chosen object was generated and spawned for both players such that there were always three objects on the table. The goal for each participant, and each session was to request and collect a total number of 36 objects. Consequently, the RGSa session ended when participants completed this task which usually took approx. 30–40 minutes. The system continuously stored the moves of both participants as well as the game events. Finally, a previously trained therapy assistant supervised all the sessions. Their role was to monitor the participants during the intervention interval and support them when a trial could not be realized independently. Critically, the assistant did not offer any elements of standard speech and language therapy. The detailed methodology of the RGSa treatment is available in [22].

2.3. Stimuli and multisensory cueing

The stimuli used in the study consisted of 120 items presented in the form of three-dimensional virtual objects (Fig. 1B). It has been widely accepted that properties of target stimuli, including visual complexity, name agreement as well as imageability, might affect picture identification and, consequently, lexical retrieval [18], [62]–[64]. Hence, to ensure visual unambiguity, all objects were first evaluated by healthy participants and clinicians involved in the trial. Furthermore, all items were categorized regarding frequency, semantic category, complexity, and phonemic similarity, and they were matched for the syllable length and semantic category.

To test our hypotheses, the one hundred twenty stimuli were classified into two categories, including (A) sixty items without semantically related such as "pancake" and (B) sixty items for which acoustic features are highly relevant such as "telephone." The first group of stimuli (i.e., A) underwent the Silent Visuomotor Cueing (SVC) [34]. The cues consisted of displaying videos that showed recordings of a speech and language therapist who articulated each of the sixty stimuli following the criteria of standard phonemic cueing. Importantly, for the purpose of this study, the voice of the recorded therapist was muted such that the cues were silent. Every video depicted a part of the face of the therapist, including mouth and nose. The videos were recorded in the Clinica de l'Hospital Universitari Joan XXIII de Tarragona,

Spain. The second group of stimuli (i.e., B) underwent the so-called Semantic Auditory Cueing (SAC). SACs consisted of providing a sound that is semantically relevant to the object selected for the request, for example, the sound of ringing for the object representing “telephone,” or the sound of an engine revving up for the object representing “car.”

For each pair of participants, the stimuli were delivered in a pseudorandomized order, counterbalanced within each week. Cueing strategies were provided to half of the practiced stimuli. Specifically, for each participant, SVCs were delivered in 50% of the items without acoustic features (group A), and SACs were delivered in 50% of the items with semantically relevant sound (group B). In both cases, the cues were provided immediately after object selection, once per trial. All participants were given a wireless headset through which they heard feedback from the system.

2.4. Measures

To evaluate the naming accuracy of the target stimuli, we administered the primary outcome measure, in particular, the Vocabulary Test (VocabT), which included all the trained items [22]. For each word, participants could score a maximum of 5 points (0: no verbal utterance, 1: utterance followed by full phonetic priming, 2: utterance followed by priming of the initial phoneme, 3: utterance followed by full silent orofacial hint, 4: utterance followed by a silent orofacial hint of the first phoneme, 5: utterance followed by no hint). The test was administered six times over the intervention period to determine the baseline (week 0), changes in accuracy at weeks 2, 4, 6, 8, as well as the follow-up period at week 16.

As the secondary outcome measure, we computed Interaction Times (ITs, see Fig. 1A) for all stimuli and therapy sessions. IT was an objective quantification of improvement in communicative effectiveness which captured the time of successful goal-oriented peer-peer interaction. Hence, we defined IT as the time interval between the selection of the target object for the request and the collection of the matching object from the opponent. Consequently, each IT included lexical access, articulation of the request, comprehension of the target word, and the motor response of the opponent. All pairs of participants remained the same during the therapy interval, which ensured that the times of motor responses were constant, thus not influencing the language-related results.

2.5. Data analysis

We used the Wilcoxon signed-rank test to evaluate within-groups changes and Mann-Whitney U-test for between-groups comparisons. All comparative analyses used two-tailed tests and a standard level of significance ($p < .05$).

3. Results

We aimed to determine the effects of SVC and SAC on naming and communication in individuals with chronic non-fluent aphasia.

First, we evaluated naming accuracy as measured by a standard clinical scale VocabT. Our results yielded significant improvement on the proposed scale from baseline at each evaluation point including week 2 (W2, $p = 0.01$), 4 (W4, $p = 0.006$), 6 (W6, $p = 0.005$), 8 (W8, $p = 0.005$), and the follow up (W16 $p = 0.005$) (see Table 1). Second, we computed the change in Interaction Times (ITs). The analysis of the evolution of the ITs throughout the intervention interval (40 days) yielded a significant decrease for all the presented stimuli including cued and non-cued stimuli ($r = -.61$, $p < 0.001$) as well as for the two subsets chosen to undergo SVC (Fig. 1B Left, $r = -.7$, $p < 0.001$) and SAC (Fig. 1B Right, $r = -.69$, $p < 0.001$), respectively (Fig. 1C Left and Fig. 1C Right). Subsequently, to estimate the effects of the two types of multisensory cues on verbal expression, we compared the ITs between cued and non-cued stimuli for all the intervention days as well as for the early and late trials (Fig. 2A). A Wilcoxon signed-rank test demonstrated a significant difference between all cued and non-cued stimuli in SVC ($p = .001$) and SAC ($p = .003$). Specifically, we found that the difference between cued and non-cued trials was statistically significant during the early therapy sessions ($N = 15$) both for SVC ($p = .002$) and SAC ($p = .001$) (Fig. 2B Upper panel). No differences in ITs were found in the late sessions for neither SVC ($p = .73$) or SAC ($p = .53$) (Fig. 2B Lower panel). Moreover, the analysis yielded no differences in ITs between non-cued SAC and SVC stimuli in the early sessions ($p = .28$) establishing that the chosen subsets did not differ regarding difficulty.

Finally, we evaluated whether changes in communicative effectiveness as measured by the ITs reflect the improvement in verbal production. To this aim, we examined the relationship between the proposed measure automatically stored by the system (IT), and naming accuracy as quantified using the clinical scale (VocabT) which showed a significant increase from baseline after the intervention (Wilcoxon signed-rank: $p = .007$) [22]. For the analysis, we extracted ITs including all cued and non-cued stimuli from all the therapy sessions for each participant and computed mean ITs collected on the date of the administration of the VocabT ± 1 day. Spearman's correlation revealed a significant relationship between the mean ITs and the VocabT scores across participants ($r = -.89$, $p = -.03$), suggesting that ITs may be regarded as a relevant measure of verbal execution in participants with chronic non-fluent aphasia.

Table 1

Outcome measures at weeks 2, 4, 6, 8, and 16 (follow-up). Bold values indicate significant differences ($p < .05$). P-values for within-group analysis were obtained with Wilcoxon signed-rank test.

Within-group analysis		
Mean (SD) - Median 95% confidence interval for the mean (lower and upper bound)		
W2	$\delta(W2-BL)$	<i>p-value</i>
86.93(10.26)-88.60	7.8(6.85)-10.73	.01
[79.19–94.67]	[2.63–12.97]	
W4	$\delta(W4 - BL)$	
90.01(9.81)-94.09	10.88(7.03)-12.54	.006
[82.61–97.41]	[5.57–16.19]	
W6	$\delta(W6 - BL)$	
92.47(10.63)-97.37	13.34(7.33)-13.19	.005
[84.45-100.49]	[7.81–18.87]	
W8	$\delta(TW8 - BL)$	
95.06(8.31)-98.44	15.93(7.56)-13.68	.005
[88.79-101.33]	[10.22–21.64]	
W16	$\delta(W16 - BL)$	
94.78(8.79)-98.77	15.65(7.02)-14.42	.005
[88.15-101.41]	[10.35–20.95]	

4. Discussion

While phonological and semantic priming has been widely established [14]–[18], to the best of our knowledge, no study has explicitly explored the effects of silent visuomotor (SVC) and semantic auditory (SAC) cues on naming in people with aphasia who display anomia. Hence, in this study, we aimed to examine the effects of the proposed multisensory priming on accuracy and communicative effectiveness for a large set of items in ten participants with stroke-induced non-fluent aphasia at the chronic stage. To this aim, we used a VR-based language-rehabilitation protocol, the RGSa [22], in which dyads of patients practiced communicative acts (i.e., making a request) in the form of a turned-based game. We administered SVCs and SACs in a pseudorandomized manner at the moment when the active player selected the object to be requested from the interlocutor. Naming accuracy for the trained stimuli was evaluated five times during the intervention and once at the follow-up period using a standard clinical scale. Moreover, the RGSa system allowed for an objective, automatic, and continuous quantification of

the priming effects on the communicative effectiveness [22]. In particular, for each participant and all the intervention sessions, we stored and computed the so-called Interaction Times (IT, Fig. 1A), which indicated the interval from object selection for the request to the reception of its matching counterpart from the opponent. This time included lexical-access, naming, and the opponent's motor response. On the one hand, we hypothesized that naming accuracy for the trained stimuli and communication effectiveness would improve through the intervention sessions of RGSa as reflected by an increase of scores on VocabT and decrease of ITs, respectively. On the other hand, we predicted that, if the proposed SVC and SAC facilitate lexical access, they may result in faster ITs as compared to the non-cued stimuli.

The vocabulary test analysis revealed that the group significantly improved naming accuracy for both cued and non-cued stimuli at each time step as compared to baseline. This finding demonstrates that the RGSa intervention had beneficial effects on the recovery of naming and, critically, the retention of the acquired changes as evidenced by the follow-up. We believe that peer-peer interactions of RGSa sessions whereby participants were required to use every day-like language might have positively influenced the frequency of communication in social situations outside of the hospital, thus reinforcing the trained vocabulary [22], [65], [66]. The analysis of ITs further supported this finding. In particular, repeated measurements statistics yielded a significant decrease of ITs over the therapy interval, which suggests a general improvement in communicative effectiveness, such that participants progressively succeeded to obtain the matching objects faster. It is noteworthy that the significant decrease of ITs occurred independently of the fact that each IT included both the verbal utterance of the active player and motor response of the interlocutor. Critically, we found that the proposed IT measure was correlated with the performance on the VocabT. This significant relationship supports the notion that ITs reflect not only the improvement in communicative effectiveness but also (implicitly) in naming accuracy captured by a clinical scale, which includes lexical access and verbal execution. This finding suggests that already at the current stage, ITs might be implemented into computer-based, wearable, or mobile technologies as (1) a therapeutic strategy that facilitates naming, and (2) a diagnostic tool for changes in word production, which does not require the assistance or supervision of a therapist. It is important to note that an alternative way to measure improvement in individuals with aphasia could consist of subtracting the time of the actual verbal utterance, in this case, the request, from the moment of the selection of the target word. We did not implement such a method in the reported study for two reasons. First, it would require a speech recognition system, which is not suitable for our sample that includes participants with moderate-to-severe stages of aphasia. Second, this method would allow for the quantification of naming speed rather than communicative effectiveness whereby the primary objective is to achieve the behavioral goal by obtaining a desired object from the interlocutor.

The central objective of this study was to determine the effects of SVC and SAC on naming and communicative effectiveness in people with aphasia. The results support our hypothesis that both SVC and SAC facilitate naming, as evidenced by the difference in ITs between cued and non-cued stimuli. This effect was significant in the early intervention days when the exposure to the target lexicon was still infrequent. Indeed, no such differences were found in the late sessions, possibly because, at this stage, the acquisition of the target stimuli reached a plateau. These findings demonstrate that both visuomotor

(SVC) and acoustic (SAC) information seems to aid learning of the trained stimuli, supporting the inference-based network perspective on naming [33].

Of fundamental clinical relevance, these results provide evidence for beneficial effects of multisensory cueing on verbal execution suggesting that integrating SVC and SAC in the rehabilitation of aphasia could foster language-production skills within and outside of the clinic and even at the chronic stages of the disease. Furthermore, the reported findings might find applications as predictors of post-stroke aphasia recovery. Specifically, there is both behavioral and neuroimaging evidence which demonstrates that responsiveness to cues (i.e., classical phonological cueing) predicts immediate treatment outcomes in other phonological treatment approaches [28], [67]. We designed SVCs such that they contain visuomotor information related to the phonology of a target word while SACs contain the auditory information related to the semantics of a target word. Thus, future studies should evaluate if responsiveness to SVC and SAC is predictive of outcomes on phonological and semantic tasks, respectively.

Of scientific relevance, our findings are consistent with sensorimotor accounts of language processing [33], [40], [68] highlighting the neuronal coupling between perceptual and motor brain regions. They also provide supporting evidence for network interpretation of speech production whereby different stages of naming are governed by principles of statistical inference across all available perceptual sources [33]. On the one hand, the beneficial effects of SVCs in the early intervention sessions might be attributed to increased activity of the language networks related to the processing of orofacial gestures, thus facilitating articulation [37], [38]. On the other hand, SACs might have facilitated word production by activating semantic regions, including pSTG and MTG, thus facilitating lexical access and consequently naming [39]. Future studies shall systematically investigate the neurophysiological underpinnings of both types of cues.

5. Conclusions

This study extends current empirical and clinical framework on language rehabilitation by showing the efficacy of multisensory cueing in fostering naming even at the chronic stages of aphasia [15], [18], [20], [69]. Critically, the proposed strategies may be easily and at a low cost integrated into digital technology that may be used after hospital discharge to improve the quality of life of the patients. Finally, our findings support the hypothesis of the inference-based network at the basis of language production [33].

Table 2
Sociodemographic patient characteristics

ID	Age	Sex	Etiology	Chronicity (m)	Severity
9	58	M	Ischemia	6	Severe
10	39	F	Ischemia	83	Severe
11	64	M	Ischemia	46	Severe
12	63	F	Hemorrhage	72	severe
13	62	M	Ischemia	106	Severe
14	56	F	Ischemia	144	Severe
15	43	F	Hemorrhage	72	Severe
16	55	F	Ischemia	6	Moderate
17	75	M	Ischemia	144	Moderate
18	61	M	Ischemia	20	Moderate
Mean (SD)	57.6 (9.9)			69.9 (48.7)	

Abbreviations

SVC- Silent Visuomotor Cueing

SAC- Semantic Auditory Cueing

ILAT- Intensive Language Action Therapy

VR- Virtual Reality

RGSa- The Rehabilitation Gaming System for aphasia

VocabT- Vocabulary Test

IT- Interaction Time

pSTG- posterior superior temporal gyrus

MTG- middle temporal gyrus

References

1. Engelter ST, et al., "Epidemiology of aphasia attributable to first ischemic stroke: Incidence, severity, fluency, etiology, and thrombolysis," *Stroke*, vol. 37, no. 6, pp. 1379–1384, Jun. 2006.

2. Goodglass H, Wingfield A. "Anomia : neuroanatomical and cognitive correlates." 1997.
3. Laine M, Martin N, *Anomia : theoretical and clinical aspects*. Psychology Press, 2006.
4. Foygel D, Dell GS. Models of Impaired Lexical Access in Speech Production. *J Mem Lang*. 2000;43(2):182–216.
5. Schwartz MF. "Theoretical analysis of word production deficits in adult aphasia," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 369, no. 1634. Royal Society, 19-Jan-2014.
6. Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. *Behav Brain Sci*. 1999;22(1):1–75.
7. Oldfield RC, "Things, Words and the Brain*," *Q. J. Exp. Psychol.*, vol. 18, no. 4, pp. 340–353, Nov. 1966.
8. Kempen G, Huijbers P, "The lexicalization process in sentence production and naming: indirect election of words," *Cognition*, vol. 14, no. 2, pp. 185–209, Sep. 1983.
9. Levelt WJM, "Spoken word production: A theory of lexical access," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 23, pp. 13464–13471, Nov. 2001.
10. Thompson CK, Jacobs B, Legrand HR. Phonological treatment of naming deficits in aphasia model-based generalization analysis. *Aphasiology*. Jan. 1993;7(1):27–53.
11. Best W, Herbert R, Hickin J, Osborne F, Howard D. Phonological and orthographic facilitation of word-retrieval in aphasia: Immediate and delayed effects. *Aphasiology*. 2002;16(1–2):151–68.
12. Nickels LA. Theoretical and methodological issues in the cognitive neuropsychology of spoken word production. *Aphasiology*. 2002;16:no. 1–2., ",, pp. 3–19.
13. Heath S, et al., "Neural mechanisms underlying the facilitation of naming in aphasia using a semantic task: an fMRI study," *BMC Neurosci.*, vol. 13, no. 1, p. 98, Dec. 2012.
14. Nickels L. "Therapy for naming disorders: Revisiting, revising, and reviewing," *Aphasiology*, vol. 16, no. 10–11. pp. 935–979, Oct-2002.
15. Wisenburn B, Mahoney K. A meta-analysis of word-finding treatments for aphasia. *Aphasiology*. 2009;23(11):1338–52.
16. van Hees S, Angwin A, McMahan K, Copland D. A comparison of semantic feature analysis and phonological components analysis for the treatment of naming impairments in aphasia. *Neuropsychol Rehabil*. Jan. 2013;23(1):102–32.
17. Lorenz A, Ziegler W. Semantic vs. word-form specific techniques in anomia treatment: A multiple single-case study. *J Neurolinguistics*. 2009;22(6):515–37.
18. Meteyard L, Bose A, "What does a cue do? Comparing phonological and semantic cues for picture naming in aphasia," *J. Speech, Lang. Hear. Res.*, vol. 61, no. 3, pp. 658–674, Mar. 2018.
19. Abad A, et al. Automatic word naming recognition for an on-line aphasia treatment system. *Comput Speech Lang*. 2013;27(6):1235–48.

20. Kurland J, Wilkins A, Stokes P. iPractice: Piloting the Effectiveness of a Tablet-Based Home Practice Program in Aphasia Treatment. *Semin Speech Lang*. Jan. 2014;35(01):051–64.
21. Palmer R, et al., “Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big CACTUS): a multicentre, single-blinded, randomised controlled trial,” *Lancet Neurol.*, vol. 18, no. 9, pp. 821–833, Sep. 2019.
22. Grechuta K, et al. Augmented Dyadic Therapy Boosts Recovery of Language Function in Patients With Nonfluent Aphasia. *Stroke*. May 2019;50(5):1270–4.
23. Kurland J, Liu A, Stokes P. Effects of a Tablet-Based Home Practice Program With Telepractice on Treatment Outcomes in Chronic Aphasia. *J Speech Lang Hear Res*. May 2018;61(5):1140.
24. Lavoie M, Macoir J, Bier N. Effectiveness of technologies in the treatment of post-stroke anomia: A systematic review. *J Commun Disord*. Jan. 2017;65:43–53.
25. Martin N, Fink R, Laine M, Ayala J, “Immediate and short-term effects of contextual priming on word retrieval in aphasia,” *Aphasiology*, vol. 18, no. 10, pp. 867–898, Oct. 2004.
26. Martin N, Fink R, Laine M. Treatment of word retrieval deficits with contextual priming. *Aphasiology*. May 2004;18:no. 5–7., “,” , pp. 457–471.
27. Madden E, Robinson R, Kendall D. Phonological Treatment Approaches for Spoken Word Production in Aphasia. *Semin Speech Lang*. Feb. 2017;38(01):062–74.
28. Nardo D, Holland R, Leff AP, Price CJ, Crinion JT, “Less is more: neural mechanisms underlying anomia treatment in chronic aphasic patients,” *Brain*, vol. 140, no. 11, pp. 3039–3054, Nov. 2017.
29. Howard D, Gatehouse C, “Distinguishing semantic and lexical word retrieval deficits in people with aphasia,” *Aphasiology*, vol. 20, no. 9–11, pp. 921–950, Nov. 2006.
30. Howard D, Hickin J, Redmond T, Clark P, Best W. Re-visiting ‘semantic facilitation’ of word retrieval for people with aphasia: Facilitation yes but semantic no. *Cortex*. 2006;42(6):946–62.
31. Davis A, Pring T, “Therapy for Word-finding Deficits: More on the Effects of Semantic and Phonological Approaches to Treatment with Dysphasic Patients,” *Neuropsychol. Rehabil.*, vol. 1, no. 2, pp. 135–145, Apr. 1991.
32. Zegarek G, Arsiwalla XD, Dalmazzo D, Verschure PFMJ. “Mapping the language connectome in healthy subjects and brain tumor patients,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016, vol. 9886 LNCS, pp. 83–90.
33. Massaro Dominic SP. *Perceiving Talking Faces: From Speech Perception to a Behavioral Principle*. 1998.
34. Grechuta K, et al., “The effects of silent visuomotor cueing on word retrieval in Broca’s aphasies: A pilot study,” in *Rehabilitation Robotics (ICORR), 2017 International Conference on*, 2017, pp. 193–199.
35. Ojemann G, Ojemann J, Lettich E, Berger M. Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. *J Neurosurg*.

- 1989;71(3):316–26.
36. Duffau H, Capelle L, Denvil D, Gatignol P, Neuroimage NS-, and undefined 2003, “The role of dominant premotor cortex in language: a study using intraoperative functional mapping in awake patients,” *Elsevier*.
 37. Petrides M, Cadoret G, Mackey S, “Orofacial somatomotor responses in the macaque monkey homologue of Broca’s area,” *Nature*, vol. 435, no. 7046, pp. 1235–1238, Jun. 2005.
 38. Nishitani N, Hari R, “Viewing lip forms: Cortical dynamics,” *Neuron*, vol. 36, no. 6, pp. 1211–1220, Dec. 2002.
 39. Kiefer M, Sim EJ, Herrnberger B, Grothe J, Hoenig K, “The sound of concepts: Four markers for a link between auditory and conceptual brain systems,” *J. Neurosci.*, vol. 28, no. 47, pp. 12224–12230, Nov. 2008.
 40. Kiefer M, Pulvermüller F. “Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions,” *Cortex*, vol. 48, no. 7. pp. 805–825, Jul-2012.
 41. Wyss R, König P, Verschure PFMJF, König P. A model of the ventral visual system based on temporal stability and local memory. *PLOS Biol.* 2006;4(5):120.
 42. Verschure PFMJ, Althaus P. A real-world rational agent: unifying old and new AI. *Cogn Sci.* 2003;27(4):561–90.
 43. Segaert K, Menenti L, Weber K, Petersson KM, Hagoort P, “Shared Syntax in Language Production and Language Comprehension—An fMRI Study,” *Cereb. Cortex*, vol. 22, no. 7, pp. 1662–1670, Jul. 2012.
 44. Rodd JM, Longe OA, Randall B, Tyler LK, “The functional organisation of the fronto-temporal language system: Evidence from syntactic and semantic ambiguity,” *Neuropsychologia*, vol. 48, no. 5, pp. 1324–1335, Apr. 2010.
 45. Acheson DJ, Hagoort P, “Stimulating the brain’s language network: Syntactic ambiguity resolution after TMS to the inferior frontal gyrus and middle temporal gyrus,” *J. Cogn. Neurosci.*, vol. 25, no. 10, pp. 1664–1677, Aug. 2013.
 46. Hickok G, “Speech Perception, Conduction Aphasia, and the Functional Neuroanatomy of Language,” in *Language and the Brain*, Elsevier, 2000, pp. 87–104.
 47. WISE R, CHOLLET F, HADAR U, FRISTON K, HOFFNER E, FRACKOWIAK R, “DISTRIBUTION OF CORTICAL NEURAL NETWORKS INVOLVED IN WORD COMPREHENSION AND WORD RETRIEVAL. ” *Brain*. Aug. 1991;114(4):1803–17.
 48. Price CJ, et al., “Hearing and saying,” *Brain*, vol. 119, no. 3, pp. 919–931, Jun. 1996.
 49. Paus T, Perry DW, Zatorre RJ, Worsley KJ, Evans AC, “Modulation of Cerebral Blood Flow in the Human Auditory Cortex During Speech: Role of Motor-to-sensory Discharges,” *Eur. J. Neurosci.*, vol. 8, no. 11, pp. 2236–2246, Nov. 1996.
 50. Töpper R, Mottaghy FM, Brüggmann M, Noth J, Huber W. Facilitation of picture naming by focal transcranial magnetic stimulation of Wernicke’s area. *Exp Brain Res.* 1998;121(4):371–8.

51. Bookheimer SY, Zeffiro TA, Blaxton T, Gaillard W, Theodore W. Regional cerebral blood flow during object naming and word reading. *Hum Brain Mapp.* Jan. 1995;3(2):93–106.
52. & Hickok U, Erhard G, Kassubek P, Helms-Tillery J, Naeve-Velguth AK, Strupp S, Strick JP. P. L. and K., “Auditory cortex participates in speech production,” *Cogn. Neurosci. Soc. Abstr.*, vol. 97, 1999.
53. Difrancesco S, Pulvermüller F, Mohr B. “Intensive language-action therapy (ILAT): The methods,” *Aphasiology*, vol. 26, no. 11, pp. 1317–1351, 2012.
54. Pulvermüller F, Mohr B, Taub E. “Constraint-Induced Aphasia Therapy: A Neuroscience-Centered Translational Method,” *Neurobiol. Lang.*, pp. 1025–1034, Jan. 2016.
55. Peña-Casanova J. “Test barcelona,” *Edicions Masson, Barcelona*, 1990.
56. Oldfield RC, “The assessment and analysis of handedness: The Edinburgh inventory,” *Neuropsychologia*, vol. 9, no. 1, pp. 97–113, Mar. 1971.
57. Folstein MF, Folstein SE, McHugh PR. ‘Mini-mental state’. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12(3):189–98.
58. Fugl Meyer AR, Jaasko L, Leyman I. The post stroke hemiplegic patient. I. A method for evaluation of physical performance. *Scand J Rehabil Med.* 1975;7(1):13–31.
59. Grechuta K, Rubio B, Duff A, Oller ED, Pulvermuller F, Verschure PFMJ. Intensive language-action therapy in virtual reality for a rehabilitation gaming system. *J Pain Manag.* 2016;9(3):243.
60. Elman RJ, Bernstein-Ellis E. The efficacy of group communication treatment in adults with chronic aphasia. *J Speech Lang Hear Res.* 1999;42(2):411–9.
61. Pulvermüller F, Berthier ML. Aphasia therapy on a neuroscience basis. *Aphasiology.* 2008;22(6):563–99.
62. Bose A, Schafer G, “Name agreement in aphasia,” *Aphasiology*, vol. 31, no. 10, pp. 1143–1165, Oct. 2017.
63. Alario FX, Ferrand L, Laganaro M, New B, Frauenfelder UH, Segui J. Predictors of picture naming speed. *Behav Res Methods Instruments Comput.* 2004;36(1):140–55.
64. Kittredge AK, Dell GS, Verkuilen J, Schwartz MF. Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cogn Neuropsychol.* 2008;25(4):463–92.
65. Berthier ML, Pulvermüller F, “Neuroscience insights improve neurorehabilitation of poststroke aphasia,” *Nat. Rev. Neurol.*, vol. 7, no. 2, pp. 86–97, Feb. 2011.
66. Stahl B, Mohr B, Dreyer FR, Lucchese G, Pulvermüller F. Using language for social interaction: communication mechanisms promote recovery from chronic non-fluent aphasia. *Cortex.* 2016;85:90–9.
67. Hickin J, Best W, Herbert R, Howard D, Osborne F, “Phonological therapy for word-finding difficulties: A re-evaluation,” *Aphasiology*, vol. 16, no. 10–11, pp. 981–999, Oct. 2002.
68. Pulvermüller F, Fadiga L. Active perception: sensorimotor circuits as a cortical basis for language. *Nat Rev Neurosci.* 2010;11(5):351.

69. Van Hees S, Angwin A, McMahon K, Copland D. A comparison of semantic feature analysis and phonological components analysis for the treatment of naming impairments in aphasia. *Neuropsychol Rehabil.* 2013;23(1):102–32.

Declarations

Ethics approval and consent to participate

This study was approved by the local ethics committee and registered on clinicaltrials.gov (identifier: NCT02928822). Before the study, all participants received a detailed description of the study, had an opportunity to ask questions and signed a written informed consent.

Consent for publication

Not applicable.

Availability of Data and Materials

The data that support the findings of this study are available from the corresponding author upon request.

Competing interests

P. Verschure (ICREA) declares to be a founder and interim CEO of Eodyne SL, which aims at bringing scientifically validated neurorehabilitation technology to society. The other authors report no conflicts. \

Contributions

K.G., B.R.B., B.M., F.P, and P.F.M.J.V. designed the study; K.G., R.E.M., R.S.S., T.U.B., and B.M.H. performed research; K.G. and P.F.M.J.V. analyzed the data; K.G., and P.F.M.J.V. wrote the paper.

Disclosures

PFMJV (ICREA) declares to be a founder and interim CEO of Eodyne S L, which aims at bringing scientifically validated neurorehabilitation technology to society. The rest of the authors have nothing to disclose.

Funding

Ministry of Economy and Competitiveness (MINECO) Retos Investigación Proyectos I+D+i Spanish Plan Nacional project SANAR and Formación de Personal Investigador (FPI) grant BES2014068791, and the European Commission, European Research Council (ERC) under agreement 341196 (CDAC).

Figures

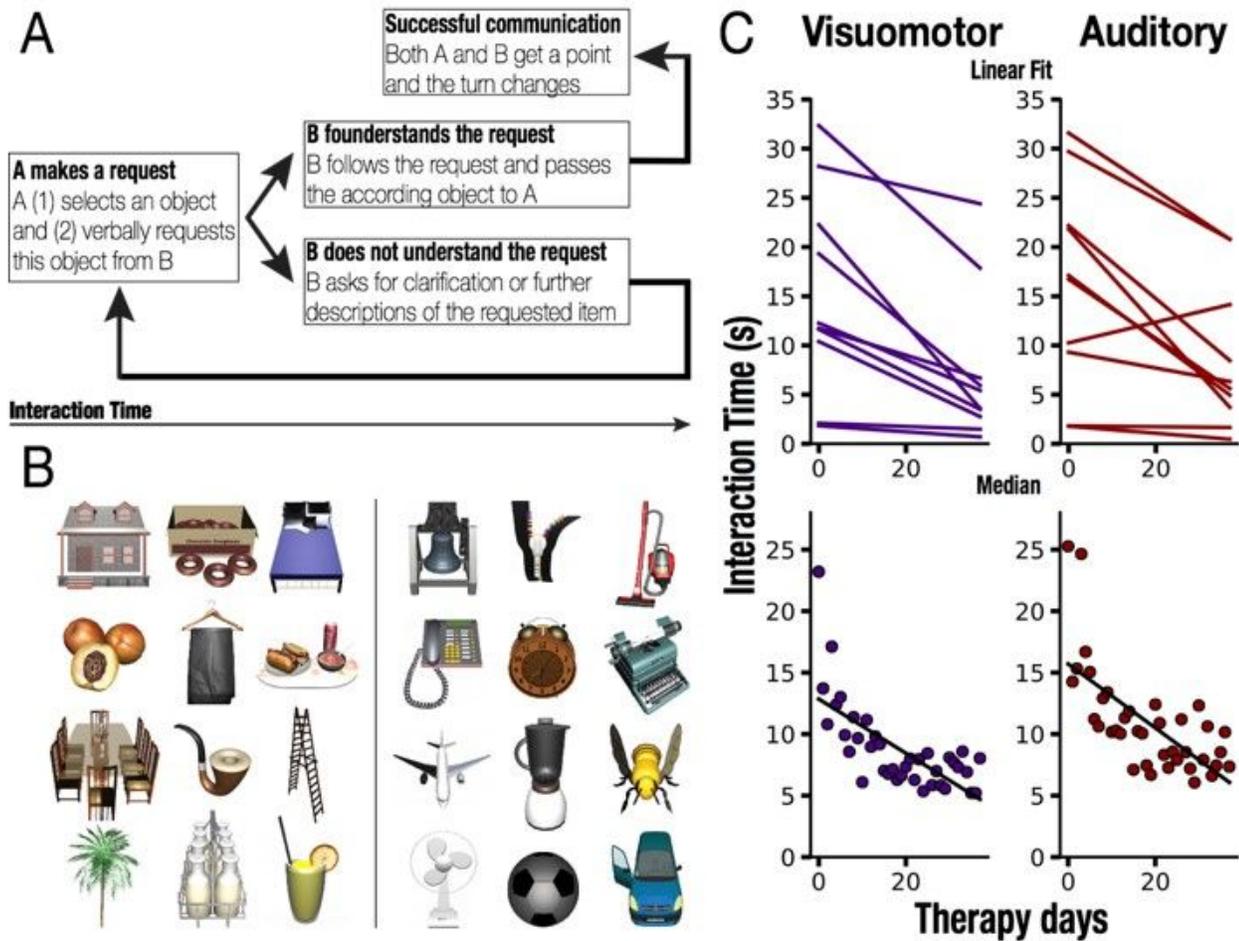


Figure 1

(A) Illustration of the Interaction Time (IT) measure, possible moves, and speech-acts. (B) Example of the materials. Left: stimuli undergoing SAC, right: stimuli undergoing SVC. (C) Fit for each participant's averaged IT over the therapy interval for all the stimuli undergoing Silent Visuomotor (SVC, violet) and Semantic Auditory (SAC, red) cueing. Upper panels: Lines represent linear regression models for individual participants including cued and non-cued trials. Lower panels: Median ITs of all the participants including all stimuli for each therapy session.

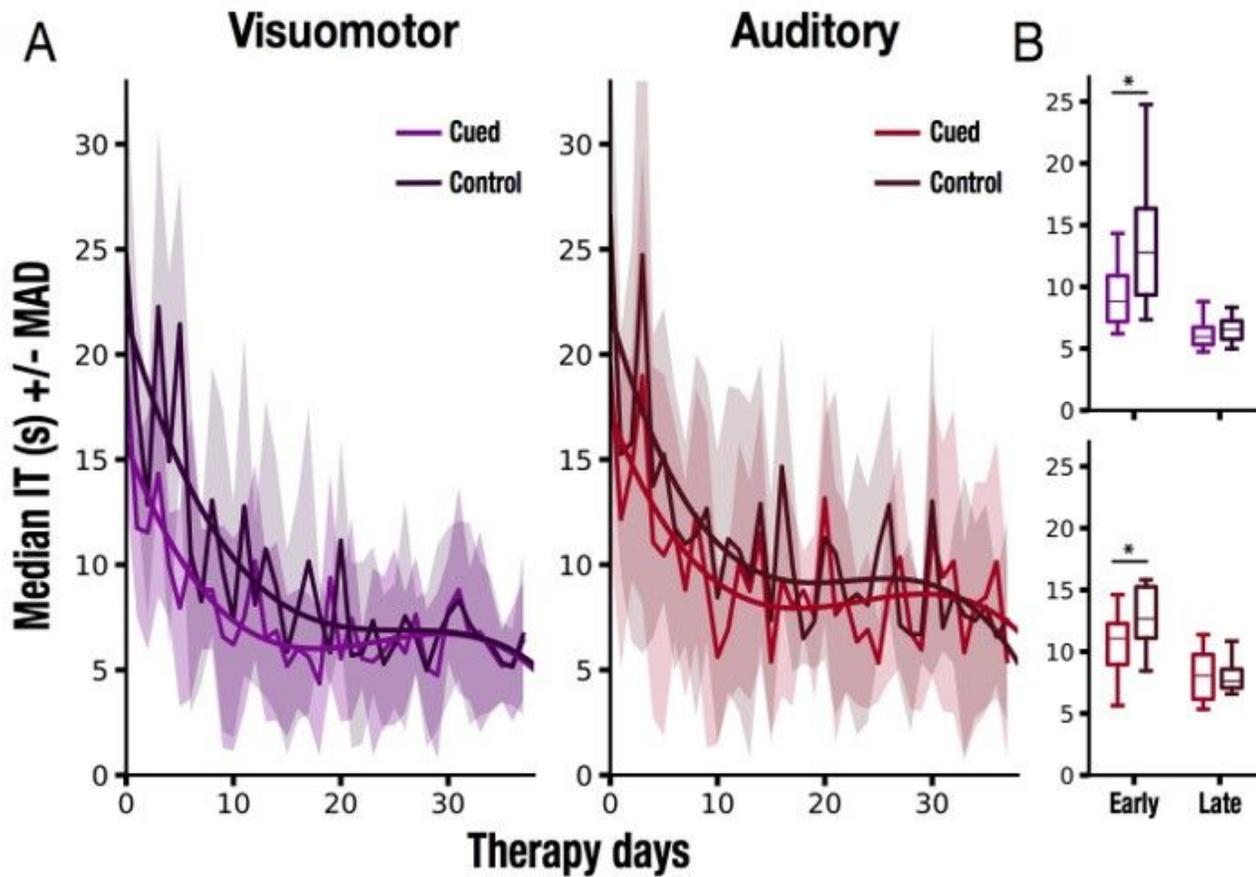


Figure 2

(A) Evolution of median ITs for cued on non-cued stimuli over the therapy sessions. Lines represent nonlinear regression models for cued and non-cued visuomotor (violet) and auditory (red) cues. (B) Quantification of differences in ITs for SVC and SAC between cued and non-cued stimuli in the early (first 15) and late (last 15) therapy sessions.