



HAL
open science

Charlie Chaplin and gesture training in severe aphasia: A controlled double-blind single-case experimental design

Clémence Vibrac, Amélie Avias, Pierre-Olivier François, Marie-Eve
Isner-Horobeti, Agata Krasny-Pacini

► To cite this version:

Clémence Vibrac, Amélie Avias, Pierre-Olivier François, Marie-Eve Isner-Horobeti, Agata Krasny-Pacini. Charlie Chaplin and gesture training in severe aphasia: A controlled double-blind single-case experimental design. *Annals of Physical and Rehabilitation Medicine*, 2021, 64 (1), pp.101356. 10.1016/j.rehab.2019.12.010 . hal-03652710

HAL Id: hal-03652710

<https://hal.science/hal-03652710>

Submitted on 26 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Charlie Chaplin and gesture training in
2 severe aphasia: a controlled double-blind
3 single-case experimental design.

4
5 **Author's copy of accepted manuscript:** Charlie Chaplin and gesture training in severe
6 aphasia: a controlled double-blind single-case experimental design

7 **Ann Phys Rehabil Med 2020 Feb 4;S1877-0657(20)30031-2.**
8 **doi: 10.1016/j.rehab.2019.12.010. Online ahead of print.**

9
10
11 **Clemence Vibrac, ST^{1,2,3}; Amelie Avias, ST³; Pierre-Olivier François, ST⁴; Marie-Eve**
12 **Isner-Horobeti, MD, PhD^{4,5}; Agata Krasny-Pacini, MD, PhD^{4,6}**

13
14 ¹ Pôle Ambroise Paré service ORL. Hôpitaux Civils de Colmar 39 avenue de la liberté 68024 Colmar cedex,
15 France

16 ² Pôle Psychiatrie service Psychiatrie Infanto Juvénile. Hôpitaux Civils de Colmar 39 avenue de la liberté 68024
17 Colmar cedex, France

18 ³ Centre de Formation Universitaire en Orthophonie de Strasbourg, 4 rue Kirschleger 67000 Strasbourg, France

19 ⁴ Pôle de Médecine Physique et de Réadaptation. Institut Universitaire de Réadaptation Clemenceau Strasbourg.
20 45 boulevard Clémenceau, F-67082 Strasbourg cedex, France

21 ⁵ Strasbourg University, Fédération de Médecine Translationnelle de Strasbourg, EA 3072 “Mitochondrie, stress
22 oxydant et protection musculaire”, France

23 ⁶ Strasbourg University, Unité Inserm 1114 Neuropsychologie Cognitive et Physiopathologie De La
24 Schizophrenie, Département De Psychiatrie, Hôpital Civil De Strasbourg, 1 Place De L'hôpital, 67091
25 Strasbourg Cedex F-67000 Strasbourg, France

26
27 **Corresponding author:**
28 **Dr. Agata Krasny-Pacini**
29 **Institut Universitaire de Réadaptation Clemenceau-Strasbourg**
30 **10 avenue Achille Baumann**
31 **BP 80096, 67402 Illkirch Cedex France**
32 **Telephone: 0033 671-28-41-52 (mobile)**
33 **E-mail: agatakrasnypacini@gmail.com**
34

36 Abstract

37 **Objective:** to explore whether two types of gesture interventions can improve communication in
38 severe aphasia. **Design:** pilot study performed at home in routine care using a controlled double-blind
39 single-case experimental design (SCED), using a multiple baseline across three subjects and across
40 two behaviors (gesture and naming). **Setting:** home-delivered intervention by an outreach team for
41 patients with acquired brain injury. **Participants:** three male patients with stroke-induced severe
42 chronic aphasia, non-functional perseverative speech and severe associated impairments.
43 **Interventions:** a passive gesture intervention in which patients watch movies selected for their
44 intensive use of gesture and an active gesture intervention in which patients actively practice gestures
45 through Visual Action Therapy. **Main Outcome Measure(s):** naming score, gesture score and non-
46 verbal subscale of the Lillois Test of Communication **Results:** Gesture interventions improved the
47 ability to gesture a list of words (Tau-U= 0.38-0.67 for combined gesture interventions effect) and
48 increased non-verbal communication activity in all three patients. Benefits were maintained at three-
49 month follow-up. **Conclusions:** mute films, that use intensive non-verbal communication may provide
50 a useful add-on to speech therapy. Improving naming in very severe and chronic aphasia may not be
51 feasible and more effort could be devoted to improving gesture-based and non-verbal communication.
52 **Key Words:** aphasia, stroke, gesture, intervention, speech therapy, single-case experimental design

53 Introduction

54 Aphasia following a stroke is a frequent and disabling condition that decreases quality of life. A
55 number of approaches have been proposed to treat aphasia, with currently no consensus¹, but the use
56 of gesture has been proposed as a way to enhance aphasia recovery². In this approach, gestures are
57 used to facilitate spoken output but also to compensate communication activity. Compensatory
58 communication relies primarily on symbolic gestures that express some type of meaning²
59 (e.g., hand shaped as a house, use of objects or actions such as a glass/drinking, familiar
60 actions such as thumbs up ... -see Sekine et al. for details³), and has shown strong
61 communicative in people with aphasia⁴⁻⁶. Therefore training gesture in aphasia may result in better
62 effectiveness of communication in everyday life (1) by enhancing the use of meaningful gestures by
63 the patients to increase the comprehension of their speech; (2) by facilitating their spoken production;
64 (3) more generally by promoting the use of non verbal communication skills. It is however unknown
65 if this may be achieved in the presence of very severe aphasia, associated with apraxia (which hinders

66 gesture production), and executive dysfunction (which hinders efficient planning of discourse and
67 inhibition of meaningless gestures and spoken productions).

68
69 Although the combination of observation and practice of gestures (e.g. Visual Action Therapy- VAT)⁶
70 seems to constitute an effective approach ⁸, the mere observation of an action (e .g. on a videoclip)
71 may enhance word production⁹.

72 The aim of the study was to explore whether two types of gesture interventions [(1) a passive gesture
73 intervention in which patients only watch movies selected for their intensive use of gesture and non-
74 verbal communication (e.g.: Charlie Chaplin) and (2) an active gesture intervention in which patients
75 actively practice gestures through VAT] can improve communication in patients with stroke-induced
76 severe chronic aphasia, through improvement of (1) meaningful gesturing ability; (2) naming ability;
77 (3) nonverbal communication skills.

78 **METHODS:**

79 The study was a pilot study performed at home in routine care, using a controlled multiple baseline
80 design across 3 subjects and 2 behaviors (gesture and naming). The first phase was a baseline without
81 specific intervention (but non-specific time with therapist), the two following phases were gesture
82 intervention phases applied in a balanced order (see table 1).

83

Table 1. Baseline and intervention description.	
Baseline	No specific aphasia training but time with therapist of same intensity (45-min sessions) as specific intervention administered in the following phases. During baseline, therapists spent time with participants, using materials not targeting aphasia but rather other cognitive functions: non-verbal logic (Logix®), mental flexibility and spatial reasoning (Tangram), visuo-spatial representation (Connect 4)
Passive gesture intervention	Watching mute films with the therapist (45-min sessions). The films were selected for the intensive use of gestures and facial expressions of their characters: mute movies (Charlie Chaplin, Buster Keaton), pantomimes (Mime Bizot), stories using sign language. In every session, different types of films were shown, to keep participants interested. Therapists did not intervene during movie screening, unless the participant asked something. At the end of the sessions, therapists asked participants questions to keep them engaged in the activity (e.g., Did you understand the movie? Are there gestures you could use? How did you like the movie?). None of the words of the repeated measures was represented in the movies.
Active gesture intervention	Visual Action Therapy, in which participants are trained to make gestures using cards and real-life objects (a telephone, a hammer, a rubber) in a progressive and structured protocol published by Helm-Estabrooks. ⁷ All items from the original Helm-Estabrooks protocol were used, as was half of the 12-word set. It was expected that the intervention would generalize to the untrained items of the 12-word set.

85

86 Phases length was determined by a restricted Marascuilo-Busk procedure¹⁰, with the restriction of
 87 having at least 5 measures per phase¹¹ i.e. phase changes of each patient were selected at random
 88 among $k=5$ possible time points, so that each phase had minimum 5 measures and maximum 10
 89 measures, with the restriction of never having the same phase length for two patients (to respect the
 90 staggered sequential introduction of intervention in SCEDs). Note that in this procedure, the number N
 91 of subjects must be less than k . The higher the difference between k and N , the more permutations ($=$
 92 $k!/(k-N)!$) the design has and the more power to detect a significant change above 0.05 probability.

93 Target behaviors were measured every other intervention session. In each phase, patients had four to
 94 five 35-45-minute-long individual intervention sessions weekly, with one of the 2 speech therapists
 95 conducting the study, in their home. Phase length was determined after 5 points of baseline were
 96 obtained. The patients were not blinded to the content of each phase but they were blinded to the

97 hypothesis of the study and to the fact that only the second and third phases had a specific content in
98 relation to their aphasia.

99

100 **Target behaviors:** At each measurement, patients had to name (target behavior 1) and gesture (target
101 behavior 2) a set of 12 words (e.g.: key, snake, glass, to dig, to color, to mix..) previously identified as
102 impossible to name (no spoken production), selected at random from a list of 100 words administered
103 prior to the study. The content of each set was patient-specific and remained the same throughout the
104 phases. Gestures were scored as follows: 0: no gesture produced; 1: gesture initiated but without
105 meaning; 2: expressive gesture but imprecise; 3: precise and expressive gesture that can be easily
106 understood. Naming was scored as follows: 0: no spoken production; 1: inappropriate spoken
107 production; 2: semantic or phonemic paraphasia 3: correct naming. Each measure being based on 12
108 words, the gesture and the naming score could vary between 0 and 36. An increase in scores meant the
109 patient was improving. The scoring was performed based on video recording of the patient, visualized
110 in randomized order. Assessors were therefore blinded to intervention phase when scoring the patient.
111 Two assessors independently scored 20 % of measures for each patient and each phase, in order to
112 calculate the inter-rater reliability of the scoring system.

113 **Other measures** included: (1) the Boston Diagnostic Aphasia Examination naming score; (2) the
114 Lillois Test of Communication (LTC)^{12,13} – non-verbal subscale (a PACE-like assessment in which
115 patient have to make the assessor discover the picture they are looking at using any form of
116 communication). LTC is a standardized assessment of functional communication, with an adequate
117 inter-rater reliability, routinely used in French-speaking aphasic patients, which assesses most aspects
118 of communication activity (motivation to communicate, body language etc...). These measurements
119 were administered four times: (1) before baseline; (2) after the first gesture intervention; (3) after the
120 second intervention; (4) three months after the end of all interventions (follow-up). This last measure
121 aimed at testing maintenance of effect and further included a supplemental measure of target behaviors
122 (12 word gesture and naming).

123 **Inclusion criteria:** patients aged 18 years and older, presenting with an ischemic left-sided stroke and
124 a chronic (onset > three years) aphasia-related severe communication disability, defined as a complete
125 inability to communicate with the physician at the medical visits, without the help of proxies.
126 Inclusion were concurrent (i.e. all patients started baseline at the same time). There were no exclusion
127 criteria.

128 The interventions and evaluation were performed at the patients' homes. The study conformed to
129 Helsinki Declaration. Patients gave a written informed consent. Procedural fidelity was measured
130 based on video-recording of the sessions by a speech therapist who did not participate in the session.

131 **Data analysis** was performed using both visual aids (using <https://manolov.shinyapps.io/Overlap/>) and
132 statistical methods adapted to single-case experimental designs : randomization tests¹⁰ based on
133 permutations for the overall effect of the interventions for all the patients and Tarlow's Tau-U¹⁴ for
134 the individual effect of the interventions for each patient. Regarding the individual effect of the
135 interventions for each patient, the following comparisons were performed: (1) baseline versus
136 cumulative effect of both gesture interventions: (2) baseline versus first gesture intervention; (3) first
137 gesture intervention versus second gesture intervention. When the number of measures at baseline is
138 seven or less, Tarlow's Tau-U method may fail to detect even high degrees of trend. For this reason, in
139 case of significant results, we checked whether Tau U effect sizes were maintained if a systemic
140 baseline correction was used for phases of length of seven or less (i.e. using a conservative and over-
141 correcting approach where even non-significant trends are corrected for). Randomization test was
142 performed using ExPRT Package¹⁵, using a within-cases comparisons of means, following a restricted
143 Marscuilo-Busk procedure¹⁰.

144 The three patients included were males, at least 4 years post middle cerebral artery ischemic stroke,
145 presenting with a right-sided hemiplegia, severe apraxia and severe communication disability due to
146 aphasia (Aphasia Severity Rating Scale of 0 in all patients) but also due to executive dysfunction. All
147 had continuous intensive on-going speech therapy since their stroke (two to four sessions weekly).
148 Other characteristics of the included patients, phase's length of each patient and interventions order are
149 presented in table 2.

150

151 **Table 2: Characteristics of patients, phase length of each patient and intervention order**

	Patient 1	Patient 2	Patient 3
Age, years	65	71	43
Time since stroke, years	6	8.5	4
Aphasia Severity Rating Scale	0	0	0
BDAE naming score pre-intervention	0	33	72
Communication	Limited to three perseverative words, including « yes /no » code understandable only to familiar persons	Correct use of yes/no. Short sentences, with mainly perseverative non-functional content (e.g. « I know everything », repeated 20 times during each medical visit)	Some naming ability but no functional communication (perseverative use of single words), good non verbal abilities (2 SD above severe aphasic patients – see figure 1)
Deficits associated to the right-sided hemiplegia, severe apraxia and severe aphasia		Seizures Behavioral and cognitive dysexecutive syndrome (with severe apathy)	Behavioral and cognitive dysexecutive syndrome (with disinhibition and impulsivity)
<u>Baseline (A)*</u>			
Content	Aspecific time with therapist	Aspecific time with therapist	Aspecific time with therapist
Phase A length	10 sessions	14 sessions	12 sessions
Number of measurement points**	5	7	6
<u>First specific intervention (B)*</u>			
Content	Passive Gesture Intervention	Active Gesture Intervention	Passive Gesture Intervention
Phase B length	12 sessions	16 sessions	18 sessions
Number of measurement points**	6	8	9
<u>Second specific intervention (C)*</u>			
Content	Active Gesture Intervention	Passive Gesture Intervention	Active Gesture Intervention
Phase C length	16 sessions	10 sessions	16 sessions
Number of measurement points**	8	5	8

*for clarity of statistical comparisons, and because gesture interventions were applied in a randomized order (and therefore differed among patients), the letters A, B, C refer to the order in which phases are compared: A refers always to baseline, B to the first specific intervention irrespective of its content and C to the second specific intervention irrespective of its content.

** Target behaviors were measured every other intervention session.

152

153

154 RESULTS

155 Procedural fidelity checklists indicated 83% correct implementation of intervention. Inter-rater
156 reliability of target behaviors was found to be good (correlation of 0.92 for gesture scores and 0.94 for
157 naming scores).

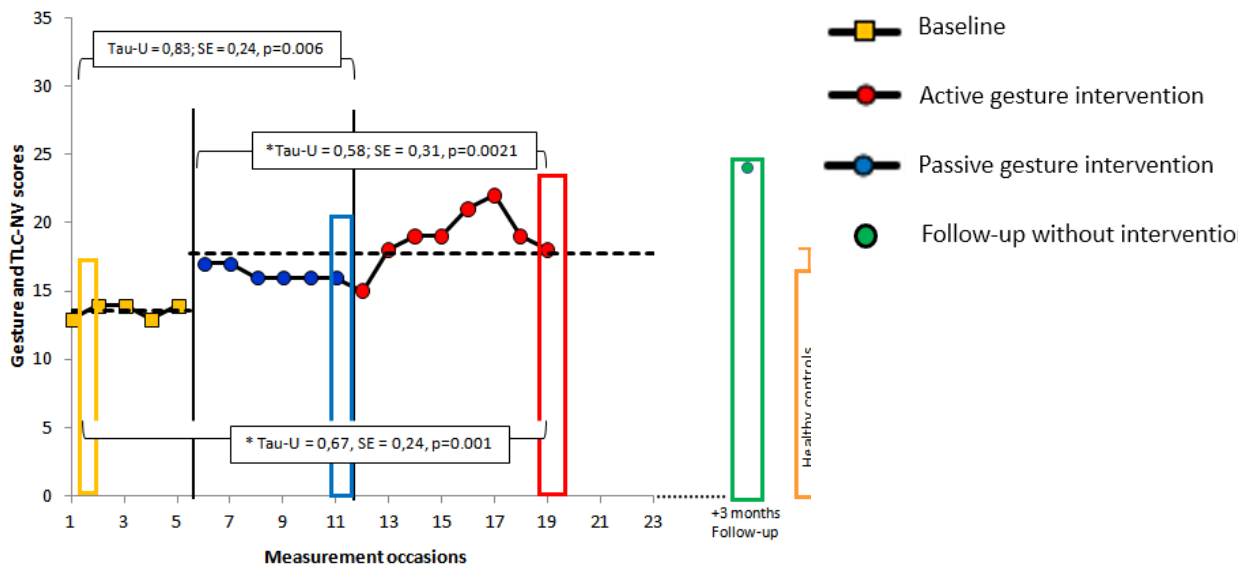
158 Gesture interventions improved the ability to gesture a list of words and increased non-verbal
159 communication activity in all three patients (see figure 1 and supplemental data 1). Active gesture
160 intervention was more effective than the passive intervention for gestures in both patients who begun
161 with the passive gesture and for naming in patient 3 (see figure 2).

162 **Figure 1: Evolution of Gesture and Non-verbal Communication scores over time, across** 163 **patients and across phases.**

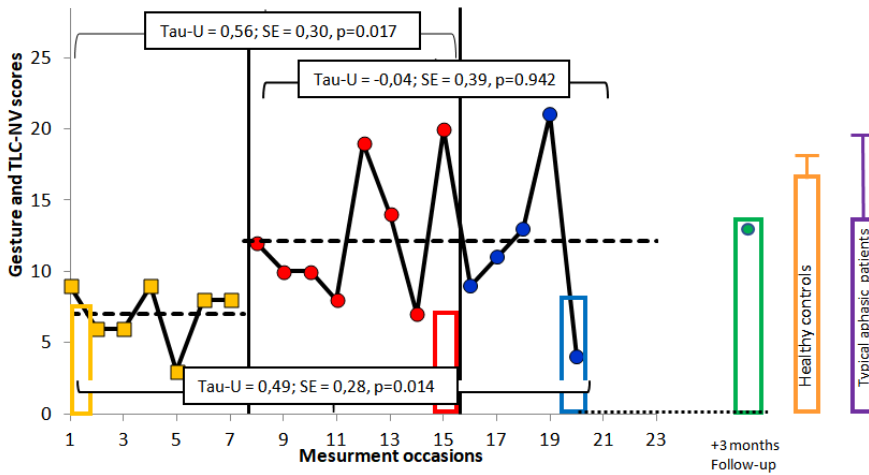
164 *Colored squares and circles represent gesture raw scores. Rectangles represent the Lillois Test of*
165 *Communication non-verbal subscale scores (TLC-NV): yellow, blue, red and green rectangles*
166 *represent the non-verbal score of the three patients included in the study respectively in baseline, after*
167 *passive gesture intervention, after active gesture intervention and at follow-up; purple rectangles*
168 *represent the typical non-verbal scores of a sample of severe aphasic patients based on Darrigrand et*
169 *al.¹³; orange rectangles represent the typical non-verbal scores of healthy controls, extracted from*
170 *LTC normative data. Error bars are 1SD. Maximum gesture score is 36 and maximum LTC-NV score*
171 *is 30. Vertical lines indicate phase changes (i.e. change in intervention content). Horizontal dotted*
172 *lines represent mean scores of baseline and mean scores of combined gesture interventions.*

173

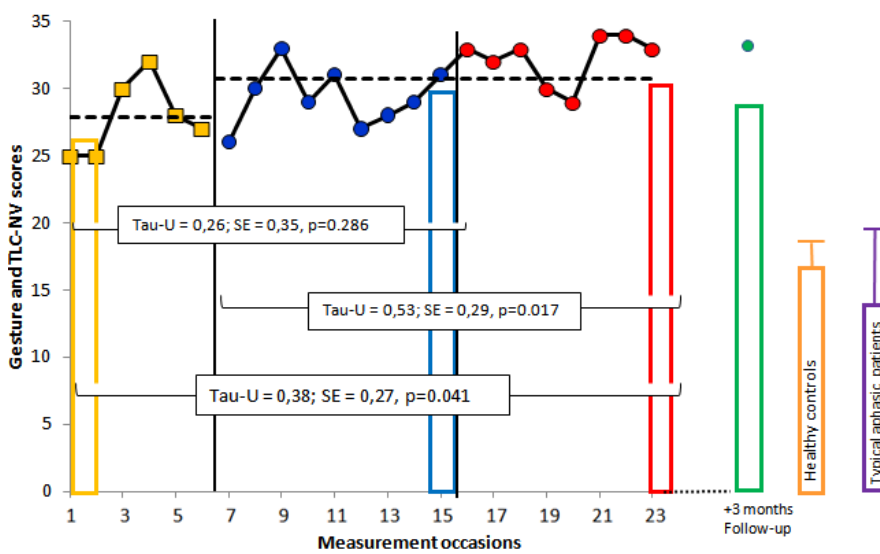
174



175



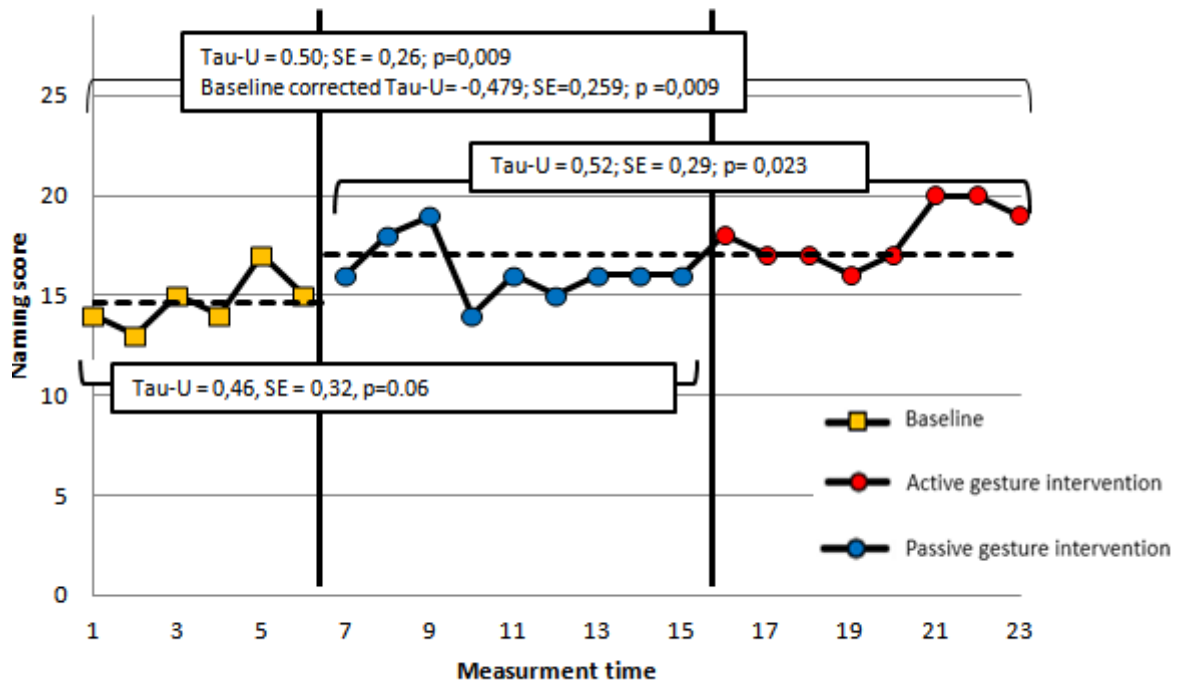
176



177

178

Figure 2: Evolution of naming score across phases for patient 3.



179

180 For patient 3, naming visual analyses (see figure 2) suggested a non-monotonic trend towards
 181 improvement of baseline, and baseline correction seemed necessary, even if the trend was not detected
 182 as significant using Tarlow's method. Using the conservative baseline correction proposed by Tarlow,
 183 overall gesture intervention (passive + active interventions taken together) suggested a statistically
 184 significant inverse effect i.e. worsening in performance (baseline corrected $\text{Tau-U} = -0.479; \text{SE} = 0.259;$
 185 $p = 0.009$). The patient did not worsen on his naming scores, as documented by increased level (i.e.
 186 mean scores of a phase, represented by the horizontal dotted line) between baseline and gesture
 187 intervention, but naming showed a more pronounced improvement during baseline than during gesture
 188 interventions. Naming did not show changes in the two other patients (see supplemental data 2).
 189 BDAE naming scores remained unchanged (0 for patient 1, 28-35 for patient 2, 66-72 for patient 3).

190

191 Overall, statistical tests using randomization tests showed that the combined effect of both gesture
 192 interventions tended to improve gesture expression abilities ($p = 0.067$), but not naming ability ($p =$
 193 0.53). Gesture interventions improved the ability to gesture a list of words and increased non-verbal
 194 communication activity (although the latter cannot be confirmed statistically due to the small number
 195 of measurement occasions, related to LTC administration time: 90 minutes). Benefits were maintained
 196 at three-month follow-up.

198 Discussion

199 This is, to our knowledge, one of the very few single-case experimental design (SCED) papers that
200 used (1) a controlled, double blind design, (2) with an adequate baseline and control for time with
201 therapist, (3) monitored treatment fidelity and (4) adequate target behavior explored for its inter-rater
202 reliability based on video recordings of patients, scored in random order. This paper meets all current
203 standards for SCEDs¹¹. Choosing phases length for a patient at random is recommended in SCEDs but
204 is very rarely done, while it considerably improves the methodological rigor of single case research.
205 The study limitation is that the Marsculio-Busk procedure may have had insufficient power ($k!(k-N)!$
206 $= 60$ possible permutation) to statistically show an intervention global effect by the statistical
207 randomization test. Higher power would have been achieved if the design allowed to randomly select
208 longer baselines (i.e. longer than 10 measures), allowing more k possible starting points and therefore
209 more permutations. It is however always challenging in SCEDs for both therapist and patients to spend
210 more time in baselines and less in intervention. A future direction for SCED research could be to
211 collect longer baselines while training the patient on another patient-relevant function, unrelated to the
212 intervention being tested. Here the baseline was a non-specific time with therapist, while it would have
213 been better to use it for relevant training (e.g.: attention, executive functioning training). Nonetheless,
214 the SCED design was strong and shows the feasibility of conducting research on small samples,
215 especially with patients who are often excluded from clinical trials because of the numerous associated
216 deficits, while they represent those most in need of innovative effective interventions.

217 A number of statistical procedures are now available to interpret SCED data, but all have their
218 limitations. The presence of improvement trends in baseline scores is a major issue, as illustrated by
219 patient 3' naming scores which showed a more pronounced improvement during baseline than during
220 passive gesture intervention. A reasonable interpretation of this result is that naming showed a practice
221 effect that probably reached a ceiling by measurement time 9 (session 18), followed by a lack of
222 further improvement with the gesture intervention. This suggests that the patient improved more with
223 the non-specific baseline intervention content, than with the gesture interventions. This raises the
224 issue of improvements unrelated to intervention such as time spent with therapist. Patient 3 was
225 young, single and quite isolated (lived in a village and was unable to drive). It is likely that spending
226 time with motivated young speech therapists played a more significant role in improving the patient's
227 scores than the content of the intervention itself. In this study, time with therapist was the same during
228 all the phases which is a methodological strength and allowed to detect this non-specific, statistically
229 significant effect.

231 Nonverbal modalities may be used as verbal facilitators/learning enhancement to improve verbal
232 communication or as a communication strategy to improve total communication¹⁶. The latter is
233 supported by our study, as gesture scores and TLC-NV scores improved, while the former was
234 supported by the results of patient 3 only for whom active gesture intervention only showed significant
235 gains in naming. Previous research has shown that gesture training alone has nonsignificant effects on
236 verbal production and should be combined with verbal training, which is probably a reason why the
237 intervention did not have the expected effect on naming².

238
239 Patients' ability to gesture words improved more with the active gesture intervention based on VAT,
240 compared to watching mute films. All 3 patients had apraxia, which has been shown to predict the
241 comprehensibility of gesturing irrespective of aphasia severity¹⁷. Apraxia may have prevented
242 patients from benefiting from the passive intervention while the repetitive step-by-step use of gestures
243 in VAT enabled learning of gestures in the presence of apraxia. This supports growing evidence for
244 the use of gesture in treating aphasia^{16,18}, use of gesture training protocols² (here VAT) and the
245 necessity of intensive treatments (here sessions weekly)¹. However to achieve better results, other
246 modalities (e.g. drawing, music...) of training should be included. On-going studies should confirm
247 that Multimodal approaches, (as opposed to Constraint-Induced Aphasia Therapy limited to spoken
248 modality) are the most effective for patients with severe aphasia¹⁸.

249 Patient 1, who had started with the passive intervention, improved his gesture scores and non-verbal
250 communication just by watching movies, though this improved further in the Visual Action Therapy.
251 Watching mute films, which use intensive non-verbal communication may constitute a useful add-on
252 to speech therapy in aphasic patients, as movies allow a more ecological use of communication (as
253 opposed to the VAT that required more artificial imitation of gestures). Generalization to spontaneous
254 discourse is a major challenge of aphasia rehabilitation². The use of films may enhance this
255 generalization because it is closer to the natural situations of communication, which is key necessity in
256 gesture treatments¹⁹. Future studies could explore the effectiveness of this approach as an add-on to
257 speech therapy

258 Improving naming in severe and chronic aphasia may not always be feasible and more effort could be
259 devoted to improving gesture-based compensatory communication and non-verbal aspects of
260 communication.

261 We have no Conflict of Interest

262

- 263 1. Brady, M. C., Kelly, H., Godwin, J., Enderby, P. & Campbell, P. Speech and language therapy for
 264 aphasia following stroke. *Cochrane Database Syst. Rev.* CD000425 (2016)
 265 doi:10.1002/14651858.CD000425.pub4.
- 266 2. Rose, M. L., Raymer, A. M., Lanyon, L. E. & Attard, M. C. A systematic review of gesture treatments
 267 for post-stroke aphasia. *Aphasiology* **27**, 1090–1127 (2013).
- 268 3. Sekine, K., Rose, M. L., Foster, A. M., Attard, M. C. & Lanyon, L. E. Gesture production patterns in
 269 aphasic discourse: In-depth description and preliminary predictions. *Aphasiology* **27**, 1031–1049
 270 (2013).
- 271 4. Rose, M. L., Mok, Z. & Sekine, K. Communicative effectiveness of pantomime gesture in people
 272 with aphasia. *Int. J. Lang. Commun. Disord.* **52**, 227–237 (2017).
- 273 5. de Beer, C. *et al.* How Much Information Do People With Aphasia Convey via Gesture? *Am. J.*
 274 *Speech Lang. Pathol.* **26**, 483–497 (2017).
- 275 6. Hogrefe, K., Ziegler, W., Weidinger, N. & Goldenberg, G. Comprehensibility and neural substrate of
 276 communicative gestures in severe aphasia. *Brain Lang.* **171**, 62–71 (2017).
- 277 7. Helm-Estabrooks, N., Fitzpatrick, P. M. & Barresi, B. Visual action therapy for global aphasia. *J.*
 278 *Speech Hear. Disord.* **47**, 385–389 (1982).
- 279 8. Ertelt, D. & Binkofski, F. Action observation as a tool for neurorehabilitation to moderate motor
 280 deficits and aphasia following stroke. *Neural Regen. Res.* **7**, 2063–2074 (2012).
- 281 9. Bonifazi, S. *et al.* Action observation as a useful approach for enhancing recovery of verb
 282 production: new evidence from aphasia. *Eur. J. Phys. Rehabil. Med.* **49**, 473–481 (2013).
- 283 10. Levin, J. R., Ferron, J. M. & Gafurov, B. S. Comparison of randomization-test procedures for
 284 single-case multiple-baseline designs. *Dev. Neurorehabilitation* **0**, 1–21 (2016).
- 285 11. Tate, R. L. *et al.* Revision of a method quality rating scale for single-case experimental designs
 286 and n-of-1 trials: the 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychol. Rehabil.*
 287 **23**, 619–638 (2013).
- 288 12. Rousseaux, M., Daveluy, W. & Kozlowski, O. Communication in conversation in stroke
 289 patients. *J. Neurol.* **257**, 1099–1107 (2010).
- 290 13. Darrigrand, B. *et al.* Communication impairment and activity limitation in stroke patients with
 291 severe aphasia. *Disabil. Rehabil.* **33**, 1169–1178 (2011).
- 292 14. Tarlow, K. R. An Improved Rank Correlation Effect Size Statistic for Single-Case Designs:
 293 Baseline Corrected Tau , An
 294 Improved Rank Correlation Effect Size Statistic for Single-Case Designs: Baseline Corrected Tau.
 295 *Behav. Modif.* **41**, 427–467 (2017).

- 296 15. Gafurov, B. S. & Levin, J. R. *ExPRT (Excel Package of Randomisation Tests): Statistical Analyses*
297 *of Single-Case intervention Data.* (2017).
- 298 16. Pierce, J. E., O'Halloran, R., Togher, L. & Rose, M. L. What Is Meant by 'Multimodal Therapy'
299 for Aphasia? *Am. J. Speech Lang. Pathol.* **28**, 706–716 (2019).
- 300 17. Hogrefe, K., Ziegler, W., Weidinger, N. & Goldenberg, G. Non-verbal communication in severe
301 aphasia: influence of aphasia, apraxia, or semantic processing? *Cortex J. Devoted Study Nerv. Syst.*
302 *Behav.* **48**, 952–962 (2012).
- 303 18. Rose, M. L. *et al.* Constraint-induced or multi-modal personalized aphasia rehabilitation
304 (COMPARE): A randomized controlled trial for stroke-related chronic aphasia. *Int. J. Stroke Off. J.*
305 *Int. Stroke Soc.* 1747493019870401 (2019) doi:10.1177/1747493019870401.
- 306 19. de Ruiter, J. P. & de Beer, C. A critical evaluation of models of gesture and speech production
307 for understanding gesture in aphasia. *Aphasiology* **27**, 1015–1030 (2013).
- 308

309 **Supplemental data 1: details of gesture scores analyses**

310 Visual analysis and overlap-based effect size: Analyzing visually the graphs, level (i.e., the
311 mean scores of each phase) improved in all participants during gesture interventions. For
312 participant 2, measurement variability increased during gesture intervention possibly because
313 of intervention intensity inducing fatigue. Participant 3 showed an important variability at
314 baseline and an important increase in scores in the first 4 measures (possibly because of a
315 practice effect) followed by a decrease in scores in the last 2 baseline points. Overlap was
316 calculated by using Tarlow's Tau-U taking into account the presence of data trends if
317 necessary (Table 2). Tarlow's method first checks for trends at baseline and then corrects the
318 data, depending on whether there are significant trends or not (baseline corrected Tau-U).

319 Tarlow's method indicated that no baseline correction was needed (non-significant
320 trend for all comparisons). Therefore, results of the first 4 columns of Table 2 are reported in
321 Figure 1. When the number of measures at baseline is ≤ 7 , Tarlow's Tau-U method may fail
322 to detect even high degrees of trend. For this reason, with significant results, we checked
323 whether Tau-U effect sizes were maintained if a systemic baseline correction was used for
324 phase length ≤ 7 (i.e., using a conservative and over-correcting approach where even non-
325 significant trends are corrected for). This is reported in Table 2.

326

327

328 Table 2. Comparison of Tarlow's Tau-U values with and without baseline correction by using
 329 <http://ktarlow.com/stats/tau/>

	Tau	SE_{Tau}	p	Baseline Corrected Tau ‡	SE_{Tau}	p
Participant 1						
A versus (BC)	0.67*	0.241	0.001	0.67*	0.241	0.001
A versus B	0.826*	0.241	0.006	0.826*	0.241	0.006
B versus C	0.581*	0.308	0.021	0.581*	0.308	0.021
Participant 2						
A versus (BC)	0.490*	0.276	0.014	0.49*	0.276	0.014
A versus B	0.564*	0.302	0.017	0.564*	0.302	0.017
B versus C	-0.036	0.392	0.942	Not computed because phase length ≥ 8 reliably rejects the need for trend correction		
Participant 3						
A versus (BC)	0.381*	0.273	0.041	-0.434*	0.266	0.019
A versus B	0.260	0.353	0.286	-0.381	0.338	0.11
B versus C	0.529*	0.291	0.017	Not computed because phase length ≥ 8 reliably rejects the need for trend correction		

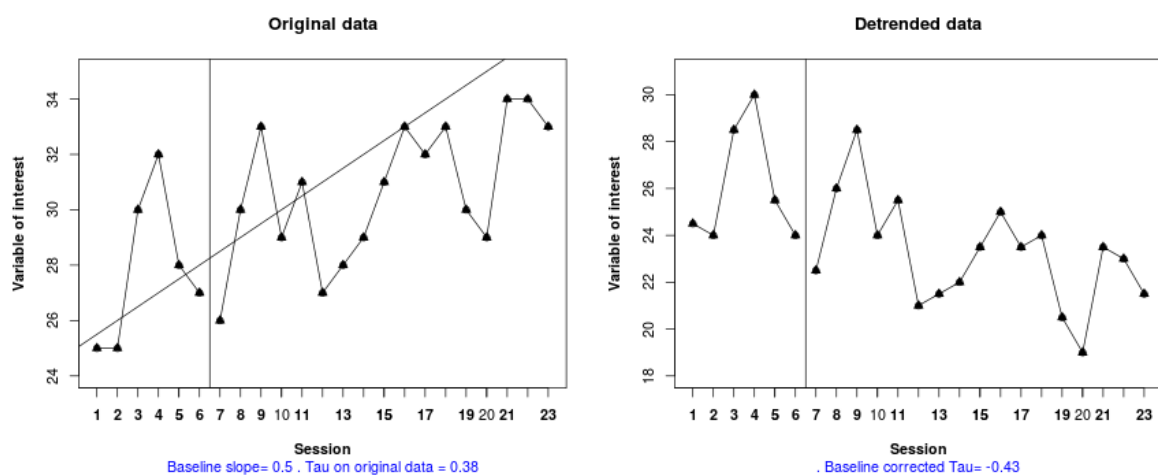
330 ‡ computed if comparing phase length ≤ 7 measures.

331 *Statistically significant at $p < 0.05$

332 Letters A, B, C refer to the order in which phases are compared: A always refers to baseline,
333 B the first specific intervention regardless of its content and C the second specific intervention
334 regardless of its content (see Table 1 and Figure 1 for the participant-specific phase B and C
335 randomized content).

336

337 Results were unchanged for participants 1 and 2. For participant 3, comparison of baseline
338 versus passive gesture intervention remained non-significant, but the effect of the overall
339 gesture intervention conferred an aberrant statistically significant worsening in performance
340 (bcTau-U = -0.434*, SE = 0.266, p = 0.019). This is not an exceptional finding and was raised
341 by Tarlow: if the projected trend line crosses the ceiling of the measurement scale, Tau-U
342 values tend to become statistically significant toward worsening (because participants cannot
343 follow a trend that goes beyond the maximum score, in this case 36). The following figure
344 illustrates the Theil-Sen regression line used in baseline corrected Tau-U (graph obtained
345 through www.manolov.shinyapps.io/overlap) and detrended data that are used for bcTau-U
346 calculation.



347

348 Figure 2. Theil-Sen regression line used in baseline-corrected Tau-U and detrended data
349 (graph obtained by using www.manolov.shinyapps.io/overlap)

350 Because bcTau-U baseline correction uses a correction of linear trend, it can lead to biased
351 baseline correction if the trend is not linear: the baseline of participant 3 was visually not
352 linear and not even monotonic; therefore, it was decided to interpret the intervention effect
353 without baseline correction in the main results of the paper.

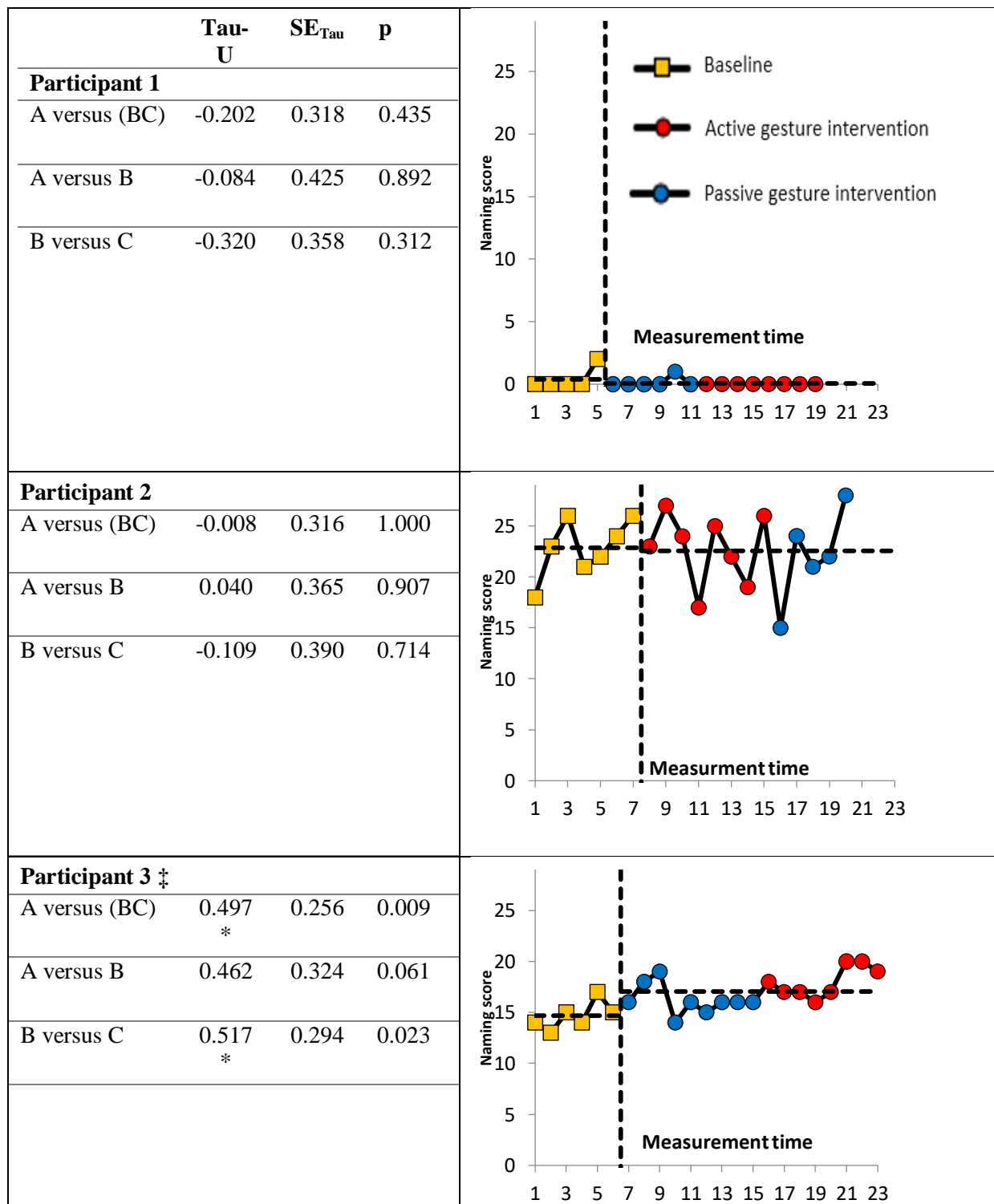
354 Another issue with Tarlow's method is that Tau-U effect sizes decrease as the number of
355 measures in intervention phase increases. This explains why for participant 1, for example,
356 Tau-U was greater when computed for A versus B (0.83) as compared with A versus the
357 cumulative effect of B and C (0.67).

358

359

360 **Supplemental data 2. Naming results.**

361



362

363

364 **Highlights**

365

- 366 • Improving naming in severe and chronic aphasia may not be feasible.
- 367 • Non-verbal modalities may be used as a communication strategy to improve total
368 communication in severe aphasia but also in some other conditions as verbal
369 facilitators to improve verbal communication.
- 370 • Mute films that use intensive non-verbal communication may be a useful add-on to
371 speech therapy.
- 372 • A controlled double-blind single-case experimental design is feasible and useful for
373 study of individuals with severe disability and associated deficits who are usually
374 excluded from large clinical trials.
- 375 • Statistical analysis must take into account the improvement trend in baselines.

376