



City Research Online

City St George's, University of London

Citation: Efstratiadou, E. A. (2018). Investigation of different therapy approaches for aphasia in the Greek language. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/19773/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

School of Health Sciences
Division of Language and Communication Science

**Investigation of different therapy approaches for
aphasia in the Greek language**

Evangelia – Antonia Efstratiadou

**Thesis submitted is presented for the Degree of
Doctor of Philosophy
of
City University of London**

Volume II

February 2018

Table of Contents

Appendices.....	3
A. Appendix A: Treatment tasks used in SFA studies.....	3
B. Appendix B: Letter of approval from Ethics Committee of University of Patras Rio Hospital.....	9
C. Appendix C: PI letter re ethics approval from Eginitio Hospital	13
D. Appendix D: Letter of approval from Ethics Committee of Division of Language and Communication Science, School of Health Sciences, City University London.....	15
E. Appendix E: Participant Information Sheet.....	16
F. Appendix F: Project consent form.....	23
G. Appendix G: Treatment Fidelity Paper	25
H. Appendix H: Supplemental Material.....	77
I. Appendix I: Chapter 5: Results - Descriptive Statistics Tables	109

Appendices

A. Appendix A: Treatment tasks used in SFA studies

Boyle & Coelho, 1995; Coelho et al, 2000; Boyle 2004; Davis & Staton, 2005; Wambaugh & Ferguson, 2007; Rider et al., 2008; Wambaugh et al., 2014; Mehta & Isaki, 2016	Peach & Reuter, 2010	Antonucci, 2009; Falconer & Antonucci, 2012	Hashimoto & Frome, 2011
Task: The investigation required from the participants with aphasia to generate semantic features in each session. Treatment Procedure: 1. The clinician asked the participant to name a target picture placed on the feature chart. <u>Original SFA chart:</u> The semantic feature were: group, association, action, properties, location, use. <u>Modifications of SFA chart:</u> - Wambaugh & Ferguson, 2007; Wambaugh et al., 2014: the semantic features were six: subject, purpose of action, related objects or actions, properties, location and how - Mehta & Isaki, 2016: the semantic features were five: use, properties,	Task: The investigation required from the participants with aphasia to generate semantic features in each session after picture scene description. Treatment Procedure: 1. The clinician presented to the participant two pictured scenes and asked to describe them, and asked two questions intended to solicit procedural information.	Task: The investigation required from the participants with aphasia to generate semantic features in each session. Treatment Procedure: The first two sessions: 1. Participants took turns naming pictures of single objects that they had demonstrated difficulty naming during baseline sessions.	Task: The investigation required from the participants with aphasia to generate semantic features in each session. Treatment Procedure: 1. The participant was asked to name the picture card either verbally or in writing when the picture was presented. <u>Modification of SFA chart:</u> the semantic features were three: group, physical – sensory properties and association

category, <u>personal memory</u>, and location. 2. Regardless of success in naming the target, the clinician guided the participant in producing its semantic features. a. To elicit features, the clinician asked questions or provided sentence completion cues, such as “:“What category does it belong to? Through prompts and questions, the clinician guided the participant to include the distinguishing features of a target to strengthen its activation by distinguishing it from similar items in the same semantic category. b. The clinician wrote the features on the chart as they were named. More than 1 word could be written in a feature box. For example, the box for physical properties typically had several entries, whereas the box for category typically had 1 entry. c. When the participant was unable to produce a feature, the clinician said it and wrote it on the chart, but only after first encouraging the participant to do the semantic processing independently. In Mehta & Isaki, 2016 study if a semantic feature was not produced, the researcher provided verbal cues and drawings to assist	2.This procedure yielded a list of failed lexical items derived from the participant’s discourse that were then used as targets for SFA treatment. After the participant had finished responding to all of the pictorial stimuli and the procedural questions, the clinician began selecting failed targets while considering whether they were an object or an action, whether they had been derived from the picture versus the procedural stimuli, or whether the participant had expressed clear awareness of difficulty with those items. Then, the clinician paired the selected target item with the stimulus for the object or action that produced the response. In this way participants were exposed to the lexical retrieval difficulty they experienced for that item; correct names were not provided at this time.	2.When participants demonstrated naming difficulty, they were prompted to provide semantic features that described the object as they were guided through a SFA chart. The chart was slightly modified from those used in previous SFA studies to include brief descriptors for each box label (e.g., Look – colour, size, shape, parts).	2. Participant was consistently encouraged to formulate verbal responses whenever possible; however, it was also allowed to write answers whenever wanted.
--	--	--	---

the participant in producing the feature; however, no cueing was provided when naming the target picture. d. If the participant said the target word as the features were being elicited, the success was acknowledged but listing of features continued until complete. e. If the participant failed to retrieve the target word even after all the features were listed, the clinician said the word, then the participant repeated it and reviewed all of its features.	3.The clinician then encouraged the participant to name the target item.	3.Open-ended questions, followed by cloze phrase cues were provided to facilitate feature production. All features were listed on the chart as they were produced by the participant. a. If participants struggled to produce a semantic feature, a phonemic cue was provided. b. If the participant failed to retrieve the target word, the feature was provided verbally for repetition when the feature was discernable. As participants became more familiar with the procedure, they began to spontaneously produce features, in which case production would be reinforced before continuing with the remaining features.	3. Second, the picture card was placed in the center of the SFA worksheet and the boxes for each of the features were filled. This step was completed regardless of whether or not participant was able to name the item. Participant was cued to state or write the feature cue for each of the boxes. One response was required for the category box, while three responses were required for the physical-sensory properties and association boxes. a. If participant was able to provide the requisite response, it was asked to write the response while the clinician verbalized the response. Although individual was able to write single words (and was asked to do so when naming the picture at the outset), responses were written for
---	---	--	--

			her occasionally because of the number of features that needed to be completed. b. If the participant was unable to provide a response, the clinician gave verbal and gestural cues to elicit the desired feature.
	4. SFA was completed, whether or not the intended lexical item was retrieved prior to or during treatment for the target item. a. If the participant said the target word in the process of listing features, the response was reinforced, but the procedure continued until complete for that item. b. If the participant was still unable to retrieve the target word following SFA, the clinician provided the name and review the previously identified features of the target.	4. In this phase, participants prompted to provide all appropriate features, even once the target name had been retrieved, to provide additional practice with the strategy. a. If a participant was unable to name the item after all semantic features had been produced, the clinician reviewed the features once more. b. If the participant still failed to name the item, a phonemic cue, followed by production of the item name for repetition, was provided to reinforce the connection between semantic features and the phonological word form.	4. When such cues were unsuccessful, the clinician provided the response both orally and in writing.

		5. Once the participant had finished, group members were encouraged to discuss whether the description provided was sufficient that they recognized the item. 6. Procedures to elicit connected speech followed a modified-PACE protocol (Davis, 2005), using new stimuli for each participant's turn. a. Participants asked to describe pictures that other group members could not see with sufficient detail that the rest of the group would be able to recognize the picture(s) being described. The process for guiding participants through SFA was similar to that outlined above except that	5. Finally, once the boxes of the SFA sheet were completed, the clinician verbally reviewed all the features. To ensure that the participant was paying attention to the appropriate written responses, participant was asked to repeat each of the written responses after the examiner. Then participant was asked to name the picture again, either verbally or in writing. a. If participant was unable to name the picture, the clinician provided the name verbally.
--	--	--	--

		participants proceeded with their description as soon as the item name was retrieved. b. Listeners were encouraged to request clarification if some part of the description was unclear. In Falconer & Antonucci, (2012) study: 7. A homework component was introduced to this study with the goals of generalising the SFA procedure to other speakers and situations beyond the research therapy sessions and to increase overall intensity of practice time.	
--	--	---	--

B. Appendix B: Letter of approval from Ethics Committee of University of Patras Rio Hospital

6^η ΥΓΕΙΟΝΟΜΙΚΗ ΠΕΡΙΦΕΡΕΙΑ
ΠΕΛΟΠΟΝΝΗΣΟΥ, ΙΟΝΙΩΝ ΝΗΣΩΝ, ΗΠΕΙΡΟΥ
& ΔΥΤΙΚΗΣ ΕΛΛΑΔΟΣ
ΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΓΕΝΙΚΟ
ΝΟΣΟΚΟΜΕΙΟ ΠΑΤΡΩΝ

Πάτρα 26.2.13
Α.Π. : 128

ΕΠΙΣΤΗΜΟΝΙΚΟ ΣΥΜΒΟΥΛΙΟ

Ταχ. Διεύθυνση : 26 500 Ρίο
Πληροφορίες : Αντωνία Γιαννίκα
Τηλέφωνο : 2610-994.721, 997.873
Φαξ : 2610-997.873

ΠΡΟΣ : Το Οικονομικό
Τμήμα

Το Επιστημονικό Συμβούλιο στην συνεδρίαση της 25.2.13 και λαμβάνοντας υπόψη την υπ' αριθμ. 42/19.2.13, απόφαση της Επιτροπής Έρευνας, Ηθικής και Δεοντολογίας, ενέκρινε το Ερευνητικό πρόγραμμα « Θαλής»: Επίπεδα Διαταραχής του λόγου Ελληνόφωνων ατόμων με αφασία : σχέσεις με ελλείμματα επεξεργασίας, εγκεφαλική βλάβη και προσεγγίσεις θεραπείας».

Το πρόγραμμα « Θαλής» χρηματοδοτεί ερευνητικά προγράμματα στα ΑΕΙ μέσω του ΕΣΠΑ. Στα πλαίσια αυτού του Προγράμματος δημιουργήθηκε μια ερευνητική ομάδα που συντονίζεται από το Εθνικό Καποδιστριακό Πανεπιστήμιο Αθηνών (Τμήμα Φιλολογίας και Τμήμα Ψυχολογίας) και το ΤΕΙ Πατρών (Τμήμα Λογοθεραπείας). Το πρόγραμμα έχει πέντε δράσεις και οι σχετικές πληροφορίες επισυνάπτονται. Στα πλαίσια του προγράμματος, εγκρίνεται η πρόσβαση σε ασθενείς με αφασία στο νοσοκομείο με σκοπό την συλλογή δεδομένων αλλά και την παροχή δωρεάν υπηρεσιών Λογοθεραπείας, σύμφωνα με το πρωτόκολλο της έρευνας. Η συλλογή δεδομένων θα γίνεται από εκπαιδευμένους ερευνητές όπως αναφέρεται στο πρωτόκολλο, υπό την ευθύνη του επιστημονικού υπευθύνου της μελέτης, στους χώρους του Νοσοκομείου ή στην οικία των ασθενών, χωρίς καμιά υποχρέωση αμοιβής για τις υπηρεσίες που θα λαμβάνουν οι ασθενείς. Η παροχή υπηρεσιών στους ασθενείς θα γίνεται κατόπιν συνεννόησης με τον επιβλέποντα ιατρό και δεν θα επιβαρύνει την θεραπεία του ασθενή. Όλοι οι συμμετέχοντες ασθενείς θα υπογράψουν σχετικό έντυπο συμμετοχής μετά από πλήρη ενημέρωση. Η χρονική διάρκεια του προγράμματος είναι τρία χρόνια, και θα δώσει την δυνατότητα στους ασθενείς με αφασία που νοσηλεύονται να έχουν δωρεάν πρόσβαση σε υπηρεσίες Λογοθεραπείας.

Επισυνάπτονται :

1. Συμπληρωμένο έντυπο επιτροπής Έρευνας, Ηθικής και Δεοντολογίας, ΠΠΓΝ Πατρών
2. Αναλυτικό πλάνο έρευνας
3. Ενημερωτικό φυλλάδιο ασθενών

4. Έντυπο Συγκατάθεσης Ασθενών

Επιστημονικός Υπεύθυνος : Αναπληρωτής Καθηγητής Τμήματος Λογοθεραπείας
ΑΤΕΙ Πατρών-Επιστημονικός Συνεργάτης Ιατρικής Σχολής
Πατρών - Παπαθανασίου Ηλίας
Ερευνητές : Δεληγιώργη Γεωργία

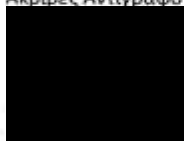
Ο ΠΡΟΕΔΡΟΣ ΤΟΥ ΕΠΙΣΤ. ΣΥΜΒΟΥΛΙΟΥ

ΚΑΘΗΓΗΤΗΣ, ΔΗΜΗΤΡΙΟΣ ΓΟΥΜΕΝΟΣ

ΤΑ ΜΕΛΗ

Αλέξ. Κουράκλη
Κων. Γιαννάκας
Αικ. Τσιατά
Βελ. Λακιώτης
Αγαρ. Κασπίρη

Ακριβές Αντίγραφο



Translated version:

**6th Health Districts
Peloponnese, Ionian Islands, Epirus
& Western Greece
University Hospital of Patras**

**Patras 26.02.2013
P. N. : 128**

Scientific Council

Address : 26500 Rio
Contact Person : Antonia Giannika
Telephone : 2610994721, 997.873
FAX : 2610997873

The Scientific Council at the meeting of the 25.02.2013 has approved the research project, "Thalis": Levels of speech disorder of Greek-speaking people with aphasia: relations with processing deficits, brain damage and treatment approaches ", with application form number No. 42/19.02.2013 to the Ethics and Research Committee.

The "Thalis" research program is funded through the NSRF. For this program a research group have been created under the coordination of the National University of Athens (Department of Literature and Psychology Department) and the TEI of Patras (Department of Speech and Language Therapy). The program has five actions and relevant information attached.

Under the "Thalis" program, researchers have authorized access to patients with aphasia in the hospital in order to collect data and provide free services of speech therapy, in accordance with the protocol of the investigation.

Data collection will be done by trained researchers as stated in the protocol, which are under the responsibility of the scientist director of the study, in the hospital or at patient's homes without any payment obligation for the services that the patients receive. The provision of services to patients would be made after consultation with the attending physician and will not disrupt the patient's treatment. All participating patients will sign a consent form after full briefing. The duration of the program is three years, and will allow patients with aphasia who are hospitalized to have free access to speech therapy.

Attached Files:

1. Completed Research Committee, Ethics Form PUGH of Patras.
2. The Research Project
3. Patient information leaflet
4. Patients Consent form

Principal Investigator: Dr. Ilias Papathanasiou, Associate Professor in Speech and Language Therapy- Speech and Language Pathology, Dept of Speech & Language Therapy, Technological Educational Institute of Patras,

The Chair of the Scientific Council

Dimitris Goumenos, Professor

Scientific Council Members

Alex. Koyrakis

Kon. Giannakenas

Aik. Tsiata

Bel. Lakiotis

Agar. Kaspiri

C. Appendix C: PI letter re ethics approval from Eginitio Hospital



NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS
SCHOOL OF PHILOSOPHY
FACULTY OF PHILOLOGY
DEPARTMENT OF LINGUISTICS

Dr Nicholas Drey
Chair Research Ethics Committee
School of Health Sciences
City University London

Athens, 2nd June 2013

Dear Dr Drey

As the Coordinating Leader and Principal Investigator of Thalís Aphasia Project, of which Eva Efstratiadou's project 'Investigation of different therapy approaches for aphasia in the Greek language' is a part, I confirm that the project has received all required ethics approvals in Greece. In particular, the project has been considered by two research ethics committees (RECs) in Greece: The University of Athens Eginitio Hospital REC, for participants recruited from the prefecture of Attica and the University Hospital of Patras, for participants recruited from Achaea. All recruiting sites in Attica accept the Eginitio Hospital REC approval.

The application form to the Eginitio Hospital REC was submitted on 09.01.2013 with protocol number 325/16-01-13. In this application we requested approval for the investigation of speech and language disorders in people with aphasia and their therapy approaches. The following documents were attached: a) A summary research protocol of the project, b) Official Declaration form that "the proposed research will be kept in strict observance of the principles of ethics and in line with what is provided in the Declaration of Helsinki. - Not used drugs", c) Official Declaration form that "There will be no charges to the Hospital when conducting the proposed research", d) Declaration form – Information and Consent form for patients.

According to the Greek Act Number 390/21.02.2013, article 18, paragraph 4 & 5 (translation attached), if within a maximum of thirty days (30) from reception of the application form,

the REC does not raise an objection to the Hospital Administration, it is considered that there is a constructive implied positive approval.

The REC of the University Hospital of Patras in their meeting on the 19/02/2013 has approved the project, which was officially announced to us with the attached letter, dated 25.02.2013 with protocol number 42/19.02.2013.

Thus, the project has met all ethics requirements in Greece and is ready to proceed. If you need any further information, please, do not hesitate to contact me.

Yours sincerely,



Dr Spyridoula Varlokosta
Associate Professor of Psycholinguistics
Department of Linguistics
Faculty of Philology
University of Athens
Greece

D. Appendix D: Letter of approval from Ethics Committee of Division of Language and Communication Science, School of Health Sciences, City University London



Research Office
Northampton Square
London EC1V 0HB

Tel: +44 (0) 20 7040 5704

www.city.ac.uk

School of Health Sciences

Ref: PhD/12-13/17

21 August 2013

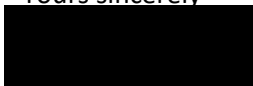
Dear Eva / Katerina

Re: Investigation of different therapy approaches for aphasia in the Greek language

Thank you for forwarding amendments and clarifications regarding your project. These have now been reviewed **and approved** by the Chair of the School Research Ethics Committee.

Please find attached, details of the full indemnity cover for your study. Under the School Research Governance guidelines you are requested to contact myself once the project has been completed, and may be asked to complete a brief progress report six months after registering the project with the School. If you have any queries please do not hesitate to contact me as below.

Yours sincerely



Alison Welton
Research Governance Officer

a.welton@city.ac.uk
020 7040 5704

E. Appendix E: Participant Information Sheet

Research Project:

Investigation of Different Therapy Approaches in Aphasia

Information Sheet for people with stroke or aphasia

You are being **invited** to take part in a **research study**. The investigation of different therapy approaches in aphasia is my **PhD project** at City University London. This study runs within the framework of Thalís Aphasia project. The Thalís Aphasia project is the largest investigation of aphasia in the Greek language and aphasia speech and language therapy (SLT) intervention. This project is conducted from the School of Philosophy in the Faculty of Philology from the Department of Linguistics at National and Kapodistrian University of Athens.

Before you decide, it is important to understand **why** we are doing this research and **what** it involves.

Please **read** this information carefully and **discuss** it with others if you wish. Please **ask** if there is **anything** that is not clear. You can talk to:

Researchers:

Evangelia – Antonia Efstratiadou
(tel: xxxx)

Ilias Papathanasiou
(tel: xxxx)

Take time to decide whether you wish to take part.

This booklet will give you information about:

- ⇒ What is the **purpose** of this **study**
- ⇒ **What we will ask you to do** if you participate
- ⇒ **What** this **project involves**

What is the purpose of the study?

You have had a **stroke**, which may have caused **aphasia**



Stroke and **aphasia** often affect the ability to express and understand language and quality of life. For relearning the abilities that have been lost or impaired **speech therapy is crucial**. In this study, we want to **investigate the effectiveness of different therapy approaches** in aphasia and which therapy has the greatest **positive effects on quality of life**.

What we will ask you to do

If you decide to take part, we will ask you:

- To give us the permission to collect information about your stroke (e.g. have access to your medical/ academic records)

Assessments

- To complete a range of **assessments** with a speech language therapist of the project.



- The speech language therapist will visit you **2-3 times** to do the following assessments:
 - A **case history** about your **stroke**

- An **aphasia** test
 - **Naming** pictures
 - A **quality of life** questionnaire
 - A **general health** questionnaire
- We will then visit you again at a convenient time a few weeks later and do the **assessments again**

Therapy

- Then we will offer you **one of two** different speech language **therapy approaches**. We will need to see you for **therapy** for **3 hours** each week, for **12 weeks**.



Re-assessment

- We will repeat the **assessments** just **after therapy** and for some of you **3 months** later.



In total, we will visit you **8-12 times** for **assessment** and **24-36 times** for **therapy**

What the project involves

If you decide to take part, we will ask you to **sign a consent form**.



If you decide to take part, you have to agree that you **will NOT receive other speech and language therapy** during the period you are involved in the project.

Assessment and re-assessment

Each visit will take **1 – 1½ hour**



If you get
another day.



tired, we will stop



and start again **later** or on

Therapy

You will have therapy either **1-to-1** with the SLT



,

Or **1-to-1** with the SLT and in a **group.**



+



When and where:

The assessments and therapies will take place at the settings where you normally receive speech and language therapy:

at your **home**



or

at your **hospital/ rehabilitation center**





The researcher will arrange a **time** to suit you for the assessment and therapy.

The **therapy approaches** will be delivered with **different ways**.



The different therapy approaches can be **individual** , **group** or both of **them together**.

Right to withdraw

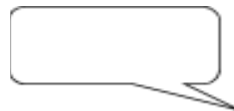
You have the right **to change your mind** about taking part at any time.



Stopping will **affect your normal care**. You will no more receive speech language **therapy**.

Confidentiality

All collected **information** will be kept **strictly confidential**



Ethics

This study has received ethics approval from the School of Community and Health Sciences of City University London (ref no: xxx).

If you are **not happy** with any aspect of the research, you need to phone or write to the Secretary of the Senate Ethics Committee. That person is: Anna Ramberg, (Anna.Ramberg.1@city.ac.uk), Academic Services, City University, Northampton Square, London EC1V 0HB, Tel: 020 7040 8010. You need to give the name of the research project as: Investigation of Different Therapy Approaches in Aphasia

What will happen to the results?

The results will be kept safely in **City University London**



We may:

- **publish** findings in journals



- **present** them at conferences



Your **name** will not be used at any time.



Benefits

There are direct benefits to **you**. This project is investigating different therapy approaches with using different therapy types. If you take part:

- You will have the opportunity to **receive a free charge, six months speech intervention program (12 weeks of therapy)**.
- Help further **stroke research** in Greece



THANK YOU FOR YOUR INTEREST IN THIS RESEARCH!!



F. Appendix F: Project consent form

CONSENT FORM FOR PARTICIPANTS

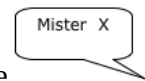
Project Title: Direct and indirect therapy approaches for aphasia in the Greek language

1. I agree to take part in the above City University research project. I have had the project explained to me, and I have read the Explanatory Statement, which I may keep for my records. I understand that agreeing to take part means that I am willing to:

- complete a range of assessments with the researcher, covering aphasia, quality of life and my general health
- participate in an therapy programme for a period of 12 weeks
- allow the researchers to have access to my medical/academic records



2. I understand that any information I provide is confidential ,
and that no information that could lead to the identification of any individual will be
disclosed in any reports on the project, or to any other party. No identifiable
personal data will be published. The identifiable data will not be shared with any other
organization.



3. Withdrawal from study

I understand that my participation is voluntary, that I can choose not to participate in

part or all of the project, and that I can withdraw  at any stage of the project.

☐☐

4. I agree to City University recording and processing this information about me. I understand that this information will be used only for the purpose(s) set out in this statement and my consent is conditional on the University complying with its duties and obligations under the Data Protection Act 1998.

☐☐

5. I agree to take part in the above study.

☐☐

Name of Participant

Signature

Date

Name of Researcher

Signature

Date

When completed, 1 copy for participant; 1 copy for researcher file.

G. Appendix G: Treatment Fidelity Paper

Treatment Integrity of Elaborated Semantic Feature Analysis Aphasia Therapy Delivered One-to-one and In-group Settings

Vasiliki Kladouchou¹, Ilias Papathanasiou², Eva A. Efstratiadou¹, Vasiliki
Christaki³, Katerina Hilari¹

¹ Division of Language and Communication Science, School of Health Sciences,
City University London

² Dept. of Speech and Language Therapy, Technological Educational Institute of
Western Greece

³ Private Practice, Athens, Greece

Corresponding author

Dr Katerina Hilari

Division of Language and Communication Science

School of Health Sciences

City University London

Northampton Square

London EC1V 0HB

UK



Tel: +44 (0) 207 040 4660

Declaration of Interest

This study evaluated the Speech and Language Therapy treatment delivered within the Thales Aphasia Project. The Thales Aphasia Project was co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: THALES – UOA - "Levels of impairment in Greek aphasia: Relationship with processing deficits, brain region, and therapeutic implications", Principal Investigator: Spyridoula Varlokosta.

Abstract

Aims: This study ran within the framework of the Thales aphasia project that investigated the effectiveness of Elaborated Semantic Feature Analysis (ESFA). We evaluated the treatment integrity (TI) of ESFA, i.e. the degree to which therapists implemented treatment as intended by the treatment protocol, in two different formats: one-to-one and group therapy.

Methods & Procedures: Based on the ESFA manual, observation of therapy videos and TI literature, we developed two ESFA integrity checklists, for one-to-one and group therapy, and used them to rate 15 therapy videos, delivered by three speech-language therapists (SLTs). Reliability of the checklists was checked, using Kappa statistics. Each sessions' TI was calculated by summing up the ratings for all components 'implemented', dividing by the total number of components 'planned' (referred to as maximum score), and then multiplying the result by 100. Differences in TI scores between the two therapy approaches were calculated, using independent sample t-tests. Treating SLTs' views on what facilitates TI were also explored through a survey.

Outcomes & Results: Inter- and intra-rater reliability were excellent ($.75 \leq \kappa \leq 1.00$) for all but one video ($\kappa=.63$). Overall, a high TI level (91.4%) was achieved. Although both approaches' TI was high, TI for one-to-one sessions was significantly higher than for group sessions (94.6% and 86.7% respectively), $t(13)=2.68$, $p=.019$. SLTs found training, use of the treatment manual, supervision, and peer support useful in implementing ESFA therapy accurately.

Conclusions & Implications: ESFA therapy as delivered in Thales is well described and therapists can implement it as intended. The high TI scores found enhance the internal validity of the main research project and facilitate its replication. The need for more emphasis on the methodological quality of TI studies is discussed.

Key Words: Elaborated Semantic Feature Analysis, Aphasia, Treatment Integrity, Treatment Fidelity, Thales Aphasia Project

What this paper adds to existing knowledge

What is already known on this subject

- Treatment Integrity (TI) is the extent to which core components of a treatment are implemented in clinical testing as intended by treatment protocols. TI data facilitate the implementation of evidence-based practice by allowing researchers to come to valid conclusions on the effectiveness of different treatments.
- Despite its importance, TI is infrequently reported: in a review of aphasia therapy studies (n=149), only 14% reported on some aspect of TI.

What this study adds

- This study provides evidence on the TI of an aphasia therapy: Elaborated Semantic Feature Analysis (ESFA). It shows that ESFA as delivered in this project was well described and therapists could effectively follow the manual and deliver the therapy as intended (TI level = 91.4%)

Clinical Implications

- Offering training, providing clinicians with a treatment manual and ongoing supervision and peer support, can help them deliver ESFA aphasia therapy as intended in order to improve the word finding difficulties of people with aphasia.
- The integrity checklists developed for this study can help clinicians monitor how closely they follow the treatment protocol.

Treatment Integrity of Elaborated Semantic Feature Analysis Aphasia Therapy Delivered One-to-one and In-group Settings

When testing the effectiveness of a treatment, like Elaborated Semantic Feature Analysis (ESFA) (Papathanasiou and Mihou 2006), it is important to ensure that the treatment is delivered by therapists as planned. Treatment fidelity refers to the methodological strategies used to monitor and enhance the reliability and validity of an intervention (Bellg et al. 2004).

Treatment fidelity was first formally defined by Moncher and Prinz (1991) who focused on delivery of treatment aspects, i.e. *treatment integrity* (TI), whether the treatment was delivered as intended, and *treatment differentiation*, whether the treatment conditions differed from one another in the intended manner (Kazdin 1986). The concept was expanded to include *treatment receipt*, which involves checking that the participant understands and can use treatment skills, and *treatment enactment*, which includes optimising the degree to which the participant is using skills learned in treatment in daily life (Lichstein et al. 1994). Further, in the Treatment Fidelity Workgroup of the National Institutes of Health Behavior Change Consortium (BCC) two more concepts were purposed: *study design*, i.e. the establishment of procedures that ensure that a study can adequately test its hypotheses, and *training provider*, which involves procedures that standardise training of therapists (Bellg et al. 2004).

The present study focuses on delivery of treatment and TI in particular. Although different terms have been used in the literature to describe TI, including procedural reliability, implementation fidelity and treatment fidelity, the term TI will be used

consistently here, defined as the extent to which core components of a treatment are implemented in clinical testing as intended by treatment protocols (Yeaton and Sechrest 1981, Dusenbury et al. 2003, McIntyre et al. 2007), in other words, therapists' adherence to the treatment protocol.

Treatment integrity has received attention in the literature because it has important implications. Firstly, TI is necessary to maintain internal validity (Moncher and Prinz 1991). It plays a key role in the interpretation of treatment results, as it allows researchers to establish whether the results of a study are attributable to the planned treatment or to the treatment that was actually implemented (Linnan and Steckler 2002, Perepletchikova and Kazdin 2005).

Treatment integrity also promotes external validity in terms of intervention replication and therefore comparisons across studies (Moncher and Prinz 1991). Treatments that can be measured for adherence to the protocol are likely to be sufficiently well described to be replicated (Mowbray et al. 2003, Hinckley and Douglas 2013). Despite the broad understanding of the importance for a study to be replicable, many studies do not meet the criteria for replication. In a literature review aiming to describe the reporting of TI data among aphasia treatment studies from 2002 to 2011, Hinckley and Douglas (2013) reported that only half of studies provided sufficient treatment description to allow replication.

The issue of TI pertains also to evidence-based practices (EBP). A critical bridge between the accumulated evidence for a treatment and its implementation in clinical practice is the understanding of its core components, which typically begins with the

establishment of integrity and the measure with which it has been assessed (Fixsen et al. 2005). Moreover, without TI data, intervention effectiveness cannot be evaluated with accuracy (Lane et al. 2004). Researchers should use a therapy protocol for training and supervising clinicians, but also for checking programme quality and performance, ensuring fidelity of the trialed intervention (Mowbray et al. 2003).

Treatment Integrity in Aphasia Therapy Studies

Despite its importance, TI is not routinely reported in speech and language therapy studies of treatment effectiveness. In aphasia therapy, which is the focus of this study, recent reviews of the literature suggest that the measurement of TI is uncommon (Cherney et al. 2008, Cherney et al. 2013, Faroqi-Shah et al. 2010, Hinckley and Douglas 2013, Rose et al. 2013). Cherney et al. (2008), focusing on constraint-induced language therapy for individuals with aphasia, included 10 studies in their review; only two of them reported data on TI. In a systematic review of 14 studies on treatment effects for bilingual individuals with aphasia (Faroqi-Shah et al. 2010), only 14% of studies checked TI. Similar findings were reported in a review of the methodological quality of 23 studies on communication partner training for people with aphasia; only 13% of them included TI data, which led the authors to conclude that one widely failed criterion across studies was TI (Cherney et al. 2013). In their review of gesture treatments for people with aphasia (PwA), Rose et al. (2013) found that 22% of the included studies reported on TI.

One could argue that the above reviews included a relatively small number of studies in specific areas and thus the notion that TI data are lacking is exaggerated. However, recently Hinckley and Douglas (2013) published the first review on the

importance of TI and the frequency with which it is reported in aphasia treatment studies. After reviewing 149 papers published between 2002 and 2011, they confirmed the results of the studies above: only 14% of studies stated clearly some aspect of TI.

Integrity Measures

In terms of methods that can be used to evaluate TI, both direct and indirect approaches exist. In direct integrity measures the researcher observes sessions, either video-recorded or live, and integrity is evaluated with the use of any sort of objective observational measure (Kaderavek and Justice 2010, Schoenwald et al. 2011). Indirect methods of integrity assessment, on the other hand, mostly include self-reports of therapists who are asked to indicate after sessions whether they included all the required components of the treatment; or self-reports of clients who are asked to report whether they received all of the components of the assigned treatment (Hinckley and Douglas 2013, Kaderavek and Justice 2010, Schoenwald et al. 2011).

In aphasia studies, the vast majority of researchers that incorporated TI measures have adopted direct methods. In particular, an independent rater checked a randomly selected sample of treatment sessions either live (Edmonds and Babb 2011, Edmonds et al. 2009, Kiran 2008, Kiran and Johnson 2008) or videotaped (Dietz et al. 2014a, Dietz et al. 2014b, Edmonds and Kiran 2006, Goff 2013, Heilemann et al. 2014, Hickey et al. 2004, Hinckley and Carr 2005, Kiran and Thompson 2003, Leonard et al. 2008, Wright et al. 2008). They used a list of core therapy components or the protocol itself to check whether each component of the treatment was implemented. To calculate adherence to the protocol, the number of components implemented by the therapists was divided by

the total number of components planned (i.e. the components that would be rated for TI) and the result multiplied by 100 (Dietz et al. 2014a, Dietz et al. 2014b, Edmonds and Babb 2011, Hickey et al. 2004). An example of a study that describes in detail the procedures followed for checking TI is that of Hickey et al. (2004).

Yet many aphasia studies that checked TI directly do not specify clearly all strategies and methods followed, for example whether they used live or videotaped observation, how they calculated the percentage of TI, or what types of scales they used to check TI (e.g. present/absent or Likert-type) (Goff 2013, Griffith et al. 2014, Kiran et al. 2011, Rider et al. 2008, Rose and Douglas 2006, Rose et al. 2002, Rose and Sussmilch 2008, Schneider and Frens 2005, Wambaugh and Wright 2007). The lack of such information creates uncertainty regarding the quality of the procedures followed and the data generated.

In terms of indirect methods for measuring TI, aphasia therapy studies have employed supervision of clinicians in conjunction with other methods, such as discussions about the treatment and its protocol as well as observations (Kempler and Goral 2011, Peach and Reuter 2010). In addition, training of providers (Goff 2013) and completion of questionnaires (Egan et al. 2004) or surveys (Heilemann et al. 2014) have also been used.

In the evaluation of TI of Semantic Feature Analysis (SFA) treatments for PwA (Boyle and Coelho 1995, Coelho et al. 2000), which is the focus of the present study, both direct and indirect methods have been adopted. The study of Peach and Reuter (2010) is an example of using indirect methods. They examined the utility of SFA for improving

verb and noun retrieval in aphasic discourse and reducing the frequency of word retrieval deficits in discourse. Their methods comprised review of the published principles for SFA therapy, discussion about them before treatment, and the presence of investigators in all treatment sessions to ensure adherence to SFA guidelines during programme implementation. In studies using direct methods to explore the TI of SFA, adherence to protocol was measured by an independent observer viewing videotaped or live sessions, as described above (Edmonds and Kiran 2006, Kiran 2008, Kiran and Johnson 2008, Kiran et al. 2011, Kiran and Thompson 2003, Rider et al. 2008).

Although a combination of several indirect measures for checking TI make integrity data more robust, these methods have low correlations with objective measures and are less reliable (Gresham et al. 2000). Direct observation is considered the gold standard in the literature as it results in more thorough and objective data. Yet, this approach also has limitations, such as staff and time requirement as well as the fact that direct observation may not represent a “natural” implementation due to the treating therapist’s awareness of observation (Cochrane and Laux 2008). Indirect data can be used to supplement objective data derived from direct methods (Heilemann et al. 2014, Hickey et al. 2004). This approach is supported by the BCC too (Bellg et al. 2004).

Research Aims

The evaluation of TI is an important part of the methodological quality of a treatment study. The present study ran within the framework of the Thales Aphasia project (<http://thales-aphasia.phil.uoa.gr>), which aimed among other factors to investigate the effectiveness of Elaborated Semantic Feature Analysis (ESFA) therapy

(Papathanasiou and Mihou, 2006), delivered through two different approaches: one-to-one therapy vs. a combination of one-to-one and group therapy. ESFA is based on the SFA approach, but also prompts the individual, after word retrieval, to elaborate the features of the word elicited on the SFA chart into a sentence. The purpose of this approach is to enable the individual to transfer the naming abilities to connected speech. We investigated the TI of the ESFA aphasia therapy in one-to-one and group therapy sessions. We focused on programme adherence, by checking therapists' consistency in the delivery of the therapy.

The research questions were:

- i What is the degree of therapists' adherence to the ESFA protocol, in one-to-one sessions, group sessions and overall in all sessions?
- ii Is there a significant difference in protocol adherence between one-to-one and group therapy sessions?

In order to facilitate the interpretation of the findings and enhance our understanding of TI, an exploration of the therapists' views on different aspects of the therapy related to TI was additionally undertaken, via an e-mail survey.

Methods

Participants

Participants in this study were the three research speech and language therapists (SLTs) who were trained in ESFA and delivered the treatment in the Thales aphasia

project. All three participants had a Master's degree, four to nine years of clinical experience and had worked with PwA from two to seven years.

People with aphasia were recruited for the main Thales aphasia project from Neurologists and SLTs working in state hospitals and private rehabilitation centers in Athens, Greece. Thirteen out of the first 16 PwA recruited were involved in this study. They had to meet the following eligibility criteria: were > 18 years old and native Greek speakers; had aphasia due to a stroke, as reported by their referring clinician; were at least four months post stroke and medically stable; had no history of other neurological or psychiatric problem and no considerable cognitive decline [scored ≥ 32 out of 38 on the Brief Cognitive Screening Test, a Greek cognitive test specifically targeted to PwA, based on items from the Dementia Rating Scale (Mattis 1988) and the Raven's Coloured Progressive Matrices (Raven 2004)]. People were excluded if they received other speech and language therapy services during the Thales project and if they did not live independently at home prior to the stroke.

Therapy videos of 13 PwA were used in this study. Aphasia was assessed with the Greek version of the Boston Diagnostic Aphasia Examination (Papathanasiou et al. 2008). Five participants had global aphasia, four Broca's aphasia, two anomic aphasia, one conduction aphasia and one transcortical motor aphasia. Of them, 61.5% were men (n=8) and the remaining 38.5% women (n=5). The participants' ages ranged from 40 to 79 years, with a mean (SD) age of 59.5 (12.1) years. Regarding their education, it ranged between 6 and 19 years [mean (SD)= 13.3 (3.8)]. In terms of their time post-stroke, PwA had a median (IQR) time post-stroke of 10 (7.0–67.5) months.

Materials and Procedures

Data and sampling procedure. All participants gave their written informed consent to take part. Within the timeframe of 10 months leading to the data analysis of this study, each of the three SLTs had to provide three one-to-one and two group therapy videos, recorded during the main research project. These videos had to meet the following criteria: the full therapy session had to be recorded, and both therapist and client(s) had to be clearly visible on the recording. The videos were recorded with a Panasonic VC-H110 video camera. They were analysed from the beginning to the end, in order for all the important components of the ESFA therapy to be checked, for each session.

Therapy Procedure. The ESFA therapy, including the stimuli selection procedure, is fully reported according to the TIDieR guidelines (Hoffmann et al. 2014) in Appendix 1.

Therapy Overview. Semantic Features Analysis (SFA) aims to improve word retrieval, by focusing on strengthening the connections between the target word and its semantic network (Boyle 2004, Boyle and Coelho 1995, Coelho et al. 2000, Conley and Coelho 2003, Lowell et al. 1995). During treatment, individuals with word retrieval difficulties are encouraged to generate words that are semantically related to the target word (i.e., semantic features), by completing a feature analysis chart. The ESFA therapy proceeds a step further, prompting the individual, after word retrieval, to elaborate the features of the word elicited on the SFA chart into a phrase and then a sentence (see one-to-one and group therapy below for more info).

In the Thales aphasia project participants were randomised to receive either 36 hours of one-to-one therapy (three one-hour sessions per week for 12 weeks) or 36 hours of a combination of one-to-one and group therapy (two 45-minute one-to-one sessions and one 1½ hour group session per week for 12 weeks). The sessions took place mainly in the participants' home and some in hospital settings.

One-to-one therapy. The therapy process is detailed in Appendix 1. In summary, during the therapy session, the client chose a picture from the stimuli set and the therapist asked them to name it. Then, presenting a semantic feature chart [same to that shown in Boyle (2004), but translated in Greek language], the therapist prompted the client to think of and say words related semantically with the target word (semantic features). The chart included 6 categories: *superordinate category*, *use*, *action*, *physical properties*, *location*, and *association*. To elicit features, the therapist asked questions or provided the client with sentence-completion cues, while prompting them to write down the features generated. If needed, the therapist used an alphabet table to help clients write; and if they were unable to write, the therapist filled in the chart.

After the chart completion and the retrieval of the word by the client, the therapist prompted the client to produce phrases with the target word and each of its features; and then to make a sentence of their choice with the target word and at least one of its features. There was no specific number of pictures to be worked on during each therapy session. The number of the pictures worked on depended on the client's abilities.

Group Therapy. During the group therapy sessions the same principles and criteria as in the one-to-one therapy were followed. The clients were asked in turn to

answer the therapist's questions to find the target word, to complete the chart, to produce phrases with therapist's cues, and finally, to produce a sentence including the target word and at least one of its features. During the process, the therapist ensured turn taking and that the same number of opportunities for response was given to each client. In addition, while during the initial therapy sessions, the therapist provided phonological or semantic cues as needed, with time, the therapist gave participants the opportunity to interact and provide appropriate cues to each other.

Integrity checklists

In order to evaluate treatment integrity of the ESFA therapy, we developed two checklists (one for individual and one for group therapy) outlining the therapy process against which TI to be checked. We also assessed reliability of the checklists to ensure their quality.

Development. The development of the ESFA integrity checklists was based on guidelines suggested by Stufflebeam (2000) and Stein et al. (2007). They were developed as a measure to be completed by an assessor, who was independent from the therapy process, but familiar with it (Heilemann et al. 2014). The checklists aimed to cover the critical therapist-oriented components of the intervention (therapists' strategies and responsibilities) in order to check their adherence to the protocol (Hogue et al. 2005).

The ESFA integrity checklists were developed by the first author who undertook two observations of live therapy sessions and had two meetings with the first author of the manual and trainer on ESFA, to ensure good understanding of the therapy and its

important components. The construction of the checklists began with the creation of a list of potential items, after identification of the primary components of the therapy, by reviewing the ESFA protocol. As the essential components of ESFA therapy - those that are expected to create therapeutic change - are not known, each therapy component that could feasibly be checked through videos and was related to therapists' responsibilities was examined (Carroll et al. 2007). Then, the potential items for the checklists were put together according to the time point of the session that they should occur.

The initial set of components was assembled as a checklist (review version) and submitted to the manual developers and therapy experts for further review and critique, in terms of relevance and comprehensiveness of the content of items, as suggested in the literature (**Netemeyer et al. 2003**). In this way, content validity of the ESFA integrity checklists was established. Based on the experts' suggestions, the checklists' content was revised, by adding, deleting or modifying the components on the list. Consensus was reached on the content and format of the checklists, through an iterative process of consultation between the developers of the ESFA therapy in this study and the authors. Two different checklists were developed this way, one for one-to-one (ESFA integrity checklist) and one for group therapy sessions (ESFA integrity checklist-G) (Supplemental Materials [insert link]).

The ESFA integrity checklists were piloted by being applied to their intended use: the first author rated four ESFA therapy session videos that were not included in the data analysis of the present study. The ratings were discussed with the last author and further changes were made to formatting and the rating method used (see below).

The final version of the ESFA Integrity checklists included three main columns labeled: (1) components, where all the relevant to TI components as described in the treatment protocol and needed to be rated for TI are listed (planned components), (2) target word, where the name of the target word worked on would be indicated, so that ratings would take place for all the words targeted during the therapy session. The inclusion of all target words for analysis was considered crucial for TI as it would allow all therapists' behaviours to be captured, which differ according to clients' performance. Moreover, adherence could be affected by the time point during the session, e.g. therapists' fatigue at the end of the session could lead to lower TI results, and (3) comments, where notes on the nature of possible deviations and troubleshooting procedures or explanation of some ambiguous ratings could be made.

Rating method of the ESFA Integrity checklists. Both Likert-type scales (Clarke 1998, Heilemann et al. 2014) and scales that capture the presence or absence of a behavior (Hinckley and Carr 2005, Schneider and Frens 2005) have been used in the literature to check TI. As TI is perceived as the *degree* to which core components of a treatment are implemented as intended, a Likert-type scale was considered the most appropriate rating method. To this end, a three-point scale was used as the rating method for the ESFA integrity checklists, where the rater was asked to use one of the following ratings: *0* (not implemented as planned), *0.5* (partly implemented as planned) and *1* (fully implemented as planned), for each component of each target word. A component could also be marked as *NA* (not applicable). Further explanation of the rating system used, with some relevant examples, is given on the checklists.

Reliability of the integrity checklists. To check inter-rater reliability of the ESFA integrity checklists, an independent rater observed and rated a randomly selected sample of three of the nine one-to-one sessions (33%) and two of the six group sessions (33%), a total percentage (33%) that is within suggested guidelines (15- 40%) (Heilemann et al. 2014). Their ratings were then compared to those of the first author, who rated all videos (n=15). For intra-rater reliability, a randomly selected sample of three of the nine one-to-one therapy (33%) and two of the six group therapy sessions (33%) were re- rated by the first author after an interval of three weeks.

E-mail survey

To facilitate the interpretation of the findings an e-mail survey was developed (Supplemental Materials [insert link]), to explore the therapists' views on different therapy aspects which are related to TI. The development of the survey's questions was based on the Implementation Fidelity Framework (Carroll et al. 2007). It aimed to explore some of the so-called *moderating factors* that may influence the degree of TI. The survey consisted of seven questions, which were categorised under three parts / possible moderating factors. The first part covered *facilitation strategies* used to support the implementation of the ESFA therapy programme. The second part elicited the therapists' views on the *ESFA manual*. As the checklists were based on the manual, therapists' views on manual properties could enhance conclusions about the face and content validity of the checklists. The third part was about *intervention complexity*, where therapists were asked to rate the complexity of ESFA therapy as low, moderate or high, based on given descriptors for the ratings. This question was added because complex interventions have

greater scope for variation in their delivery (Carroll et al. 2007), and therefore some components may be more likely not to be implemented as they should.

Data analysis

For the calculation of inter- and intra-rater reliability of the integrity checklists, Kappa statistics were used (the Kappa coefficient of Cohen) (Cohen 1960). A Kappa coefficient of .75 - 1.00 is excellent, .60 - .74 is good, .40 -.59 is fair, and below .40 is poor (Cicchetti and Sparrow 1981; as cited in Cicchetti 1994, p. 286). These guidelines are in line with benchmarks that have suggested a level of 70% and above to be regarded as an acceptable level of agreement (Heilemann et al. 2014). For TI, the first authors' ratings were used in the analyses. The TI score for each session was calculated by summing up the ratings for all the components 'implemented' (components rated as 0.5 and 1) and dividing the results by the total number of the applicable components 'planned' (referred to as maximum score). All scores were converted to percentage scores for comparability. The TI scores for (a) one-to-one sessions (n=9), (b) group sessions (n=6) and (c) overall (n=15) were calculated by summing up the scores for each session and dividing by the number of sessions (n). Different authors have considered different degrees of integrity as high (Carroll et al. 2007, Clarke 1998). However, because the level of TI that can be 'tolerated' in clinical implementation is not yet known (Kaderavek and Justice 2010), for the purposes of the present study the benchmarks suggested by Heilemann et al. (2014), which are based on a literature review, were adopted. Thus, a percentage of 80% and above was accepted as a high level of TI. Differences between one-to-one vs group sessions on adherence percentage were analysed with an independent

samples t-test, as data were normally distributed (Shapiro Wilks $p = .115$). All analyses were carried out on IBM SPSS v.22.

Results

Integrity Checklists' Reliability

Inter-rater reliability. Table 1 details the inter-rater reliability values separately for each of the five sessions, including TI scores given by the two raters. There was an excellent level of agreement between the two raters for all videos, apart from video 3 where the agreement was good (.63). The average Kappa was .82 ($p < .001$), indicating an excellent agreement between the two observers' ratings.

[table 1 about here]

Intra-rater reliability. Table 2 presents intra-rater reliability values separately for each of the five sessions selected, including TI scores given by rater 1 at two different time points. There was an excellent level of agreement between time 1 and time 2 (three weeks later) ratings, for all sessions. The average Kappa across the five sessions was .98 ($p < .001$), indicating excellent intra-rater reliability.

[table 2 about here]

Treatment Integrity (TI)

Fifteen videos of ESFA therapy sessions were rated using the ESFA integrity checklists, in order to examine the degree to which therapists adhered to the ESFA

protocol (TI score), in one-to-one sessions (n=9), group sessions (n=6), and overall in all sessions.

Treatment integrity for one-to-one therapy approach. Table 3 details the number of components planned and implemented per session and the TI scores. The overall number of planned components across the sample of the nine individual therapy sessions was 450, with the number of components per session varying as the number of target words presented to each session was dependent on the clients' performance. The mean number of components planned per session was 50 (SD=16.5) with a range between 21 and 69. Concerning the components implemented by the therapists, they were 424 (out of 450) in total, ranging across sessions from 21 to 67, with a mean (SD) of 47.3 (15.8). In terms of the session-specific TI scores for the one-to-one therapy approach they ranged 87% - 100%, with a mean TI score across all sessions of 94.6% (SD=4.6), showing a high level of TI.

[table 3 about here]

Treatment integrity for group therapy approach. Table 4 details the number of components planned and implemented per group therapy session and the TI scores. Across the six group therapy sessions, the overall number of planned components was 386, with the number of components per session varying as the number of target words presented to each session was dependent on the clients' performance. The mean number of components planned per session was 64.3 (SD=25.9), with a range between 28 and 98. Concerning the components implemented by the therapists, they were 334 in total, ranging across sessions from 25 to 77 [mean (SD) = 55.7 (22.3)]. In terms of the session-

specific TI scores for the group therapy, they ranged from 77.2% to 92.6%, with a mean (SD) TI score of 86.7% (6.9). This shows a high level of TI.

[table 4 about here]

Overall TI score. As can be seen from the results of the ratings of the one-to-one therapy and group therapy sessions, the components planned for the whole sample of 15 sessions were 836, while the components implemented by the therapists were 758, representing an overall (SD) TI score of 91.4% (6.7), with scores ranging from 77.2% to 100%. Relating the TI scores to the cut-off value of 80%: 13 of the 15 videos had TI scores > 80%, with 10 of them > 90%; two sessions (videos: 10, 11) had TI scores below 80% (78.6% and 77.2% respectively).

In summary, therapists showed a high level of TI for one-to-one therapy sessions (94.6%), for group therapy sessions (86.7%) and overall for all therapy sessions (91.4%).

Difference in treatment integrity between one-to-one and group therapy sessions. The TI scores of all one-to-one sessions (n=9) were compared with the TI scores of all group therapy sessions (n=6). The TI score for group therapy was significantly lower [mean (SD)=86.7% (6.9)] than the one for one-to-one therapy [mean (SD) = 94.6% (4.6)], ($t(13) = 2.68$, $p = .019$).

Treating SLT views - Survey

Facilitation Strategies. The first part of the survey was related to facilitation strategies used to support the implementation of the ESFA therapy programme (see figure

1). Therapists indicated training, use of the treatment manual, supervision and support by developers, and peer support as useful strategies to facilitate an accurate implementation of the ESFA programme. One therapist also found role-playing useful. On average, they rated these strategies as being of a very good to excellent quality.

[figure 1 about here]

ESFA Manual. All respondents found the ESFA manual adequate in terms of its content and rated its properties, including ease of use, clarity and comprehensiveness, as very good to excellent (see figure 2). When therapists were asked if there were any therapy components included in the manual that the therapist should be more flexible on how to implement, rather than just following the manual, responses varied. While one of the SLTs believed that a therapist should be flexible with the manual in some cases, the other two indicated that the manual instructions should be followed without deviations. The former justified her opinion by stating that not all clients are able to strictly follow the manual's instructions and thus some therapy components should be adjusted to suit clients' strengths and weaknesses. All three participants rated their adherence to the treatment manual as high (4; on a scale 1-5).

[figure 2 about here]

Treatment complexity. One of the therapists (therapist 1) indicated that ESFA therapy has a high-level of complexity (all complexity dimensions applied), while therapists 2 and 3 found ESFA moderately complex (some of the complexity dimensions applied), as detailed in Table 5.

[table 5 about here]

Discussion

Integrity Checklists' Reliability

Both TI checklists developed for this study had high inter- ($\kappa = .82$) and intra-rater ($\kappa = .98$) reliability, suggesting they are reliable measures for checking the therapists' adherence to the ESFA protocol and stable measures for TI evaluation. Other aphasia researchers who checked the reliability of their TI tools using statistical coefficients had similar findings. Heilemann et al. (2013) for instance, who used a tool similar to this study to examine TI, tested inter-rater reliability with an intra-class correlation coefficient (ICC) and found an excellent level of agreement between the two raters for all but one session, where ICC was fair (ICC = .57). The small sample of videos included in the inter-rater investigation (n=3), however, should be kept in mind when interpreting these reliability results.

Other TI studies have reported point-to-point agreement as an inter-rater reliability measure (Yoder and Symons 2010) and found a high level of agreement too (96% and above) (Kiran and Johnson 2008, Leonard et al. 2008, Wright et al. 2008). These findings however should be interpreted with caution, as they are likely to be inflated due to the fact that percentages of agreement do not correct for agreement expected by chance – a reason why reporting percentages of agreement, without including statistical coefficients, has received criticism as a measure for inter-rater reliability (Hallgren 2012).

It is worth noting that the majority of aphasia studies that adopted direct methods to examine TI do not provide evidence of inter-rater reliability of their instruments. Most employed only one independent rater to check for adherence to the therapy protocols (Dietz et al. 2014a, Dietz et al. 2014b, Edmonds and Babb 2011, Edmonds and Kiran 2006, Edmonds et al. 2009, Goff 2013, Griffith et al. 2014, Hickey et al. 2004, Kiran and Thompson, 2003, Kiran 2008, Kiran et al. 2011, Rider et al. 2008, Rose and Douglas 2006, Rose et al. 2002, Rose and Sussmilch 2008, Schneider and Frens 2005, Wambaugh and Wright 2007). Furthermore, no aphasia TI studies were found that checked intra-rater reliability; some researchers set it as a future goal though (Heilemann et al. 2013). Given that the evaluation of TI is dependent on the psychometric soundness of the TI tool used, the lack of information about reliability or the use of inadequate methods for checking it (such as point-to-point agreement) creates uncertainty for the tools that have been used in some aphasia studies and in turn for the TI scores reported. This constitutes a gap in the TI literature that should be addressed further in future research.

Some attention should be given to video 3 in the current study, which showed a lower level of agreement between rater 1 and 2, with a good but not excellent inter-rater reliability ($\kappa = .63$). A closer look on the ratings indicated a systematic pattern in the non-agreed components: the majority of differences between raters regarded the type of paraphasia produced by the client; rater 1 considered most of the client's paraphasias as circumlocutions, while rater 2 tended to consider them as semantic. The specific client was able to produce two- or three-word phrases/ structures, e.g., "turn-on, turn-off, button" for the target word 'light switch'. While for rater 1 such productions were a trial

for description of the target word (circumlocution), rater 2 considered them as semantic paraphasias mainly because of the brevity of the responses. This pattern implies that more specific rating instructions for such cases are probably needed.

Treatment Integrity

TI degree across the therapy sessions observed was high for one-to-one (94.6%), group (86.7%), and overall all sessions (91.4%). This illustrates that the therapists implemented components of the ESFA therapy as intended by the treatment protocol with high integrity. These results were consistent with SLT participants' survey replies, as all of them indicated that they implemented the therapy with a high level of integrity (4/5).

Similar findings have been reported in other aphasia studies. In particular, the majority have reported a high TI score (92% and above) (Dietz et al. 2014a, Dietz et al. 2014b, Edmonds and Babb 2011, Edmonds et al. 2009, Griffith et al. 2014, Heilemann et al. 2014, Hickey et al. 2004, J. Hinckley and Carr 2005, Rider et al. 2008, Rose and Douglas 2006, Rose et al. 2002, Rose and Sussmilch 2008, Wambaugh and Wright 2007). When more than one rater was used for checking therapists' adherence, TI score was reported in the form of point-to-point agreement between the raters, and was high as well, varying between 96%-100% (Kiran and Johnson 2008, Leonard et al. 2008, Wright et al. 2008).

Facilitation strategies used to enhance therapy implementation have probably contributed to the high TI scores found in the present study. According to SLTs' replies to the e-mail survey, training, use of the treatment manual, supervision and support by

developers and peer support were used to facilitate an accurate implementation of the ESFA programme. Such strategies have been found to optimize and standardize TI: ‘the more that is done to help implementation, through monitoring, feedback, and training, the higher the potential level of implementation fidelity achieved’ (Carroll et al. 2007, p. 7). In addition, the fact that therapists characterized the manual properties (ease of use, clarity and comprehensiveness) as very good to excellent, could have also optimized therapy implementation. Specificity enhances adherence and the comprehensiveness of a therapy’s nature can influence how far the therapists successfully adhere to its prescribed components when implemented (Carroll et al. 2007).

Two of the videos (10 & 11) scored below 80% (78.6% & 77.2% respectively), showing lower adherence. Both these videos were group therapy sessions, which by nature required more therapy components to be implemented by the therapists, making the treatment more complex and therefore more susceptible to variation in its application compared to the one-to-one approach. Moreover, both these sessions were carried out by the same SLT participant (therapist 1). Therapist 1 was the only SLT who felt in the survey that the therapist could deviate from the manual. She found that not all clients are able to strictly follow the manual’s instructions and thus some therapy components should be adjusted to suit client’s strengths and weaknesses. This shows that according to the SLT’s views therapist’s drift is justifiable. Therapist drift refers to the modification of a treatment protocol in small and gradual ways, unintentionally or unknowingly, so that a clinician amends the original protocol in an attempt to respond to a client’s specific needs and behaviors (Hinckley and Douglas 2013). Although therapist-drift threatens TI, it is

acknowledged that there is a conflict in situations where a therapist feels the obligation to comply with the protocol, but at the same time believes that a deviation from the prescribed treatment would be more helpful to their clients, and thus faces ethical and legal dilemmas (Aradi and Piercy 1985, Sweifach and Linzer 2015). In such instances, the therapist's belief in conjunction with the fact that 'trialists may struggle to exchange their role of providers of individualised care with that of researchers required to follow standardised trial procedures' (Lawton et al. 2011, p. 7) makes this therapist more prone to deviations from the manual.

Treatment integrity scores for group therapy sessions were significantly lower [mean (SD) = 86.7% (6.9)] than for one-to-one therapy sessions [mean (SD) = 94.6% (4.6)], ($t(13) = 2.68$, $p = .019$). This finding is not surprising. Findings from other fields, e.g. mental health, are in line with this, as protocol adherence was significantly higher in individual than in group therapy sessions (Long et al. 2010). It is reasonable to expect a treatment protocol to be easier to follow when there is only one client in a session. Moreover, ESFA group therapy is more complex in nature, as more therapists' behaviours are anticipated, such as prompts for interaction between the clients and turn-taking control, and it includes more interacting and interconnecting components (Craig et al. 2008); this was also evident by SLT participants' survey replies in terms of therapy complexity. Both of these factors could explain group therapy's lower TI, as it is easier to reach high integrity in simple than complex interventions (Dusenbury et al. 2003). Trying to indicate possible sources of heterogeneity in implementation of group ESFA therapy and address them in a next step could be a useful strategy for achieving even

higher TI scores for this approach (Carroll et al. 2007).

Limitations and Directions for Future Research

Limitations of the study include the checklists comprising all rather than only active ingredients of ESFA therapy, and the video sampling method. The active ingredients of an intervention distinguish essential from non-essential therapy components (Abry et al. 2015). When the active ingredients of a therapy, like ESFA, are not known all therapy components need to be examined (Carroll et al. 2007), an approach that was followed in this present study. Yet, establishment of the active ingredients of ESFA therapy would facilitate the identification of the relative importance of each component, which is crucial when guidelines for evaluating integrity are developed (Gresham et al. 2000) and the creation of meaningful thresholds of TI for the ESFA therapy. It can also provide guidance to practitioners on what to prioritise to get the most leverage from the therapy. Sensitivity or component analysis needs to be conducted using TI information and performance outcomes from a number of ESFA therapy studies to determine which components or combination of them are essential (*i.e.*, they are prerequisite if the therapy is to have its expected effect) (Carroll et al. 2007). When the active ingredients will have been identified, the current TI checklists can be modified into more precise tools.

In this project, each of the three SLTs had to record a specific number of videos meeting specific criteria, and these videos were used for the TI evaluation. However, a randomly selected sample should preferably be analysed (Heilemann et al. 2014) in order for the videos sample to be as representative as possible and eliminate sampling bias. Due

to technical difficulties, this was not feasible for the present study and should be kept in mind when interpreting the findings.

Evaluation of other aspects of TI such as *quality of delivery* and *participant responsiveness* could be targeted in future TI studies of the ESFA therapy, as the degree to which full adherence is achieved may be moderated by these two factors (Carroll et al. 2007). *Quality of delivery* refers to the manner in which a provider delivers a programme, while *participant responsiveness* focuses on the clients, and measures how far they respond to, or they are involved in a therapy, including their judgments about the outcomes and relevance of an intervention (Carroll et al. 2007). Kaderavek and Justice (2010) recommend that quality delivery evaluation is important as ‘a treatment can be implemented badly even when adherence to the procedure is high’ (p. 372). To this end, it is important first to explore which therapist skills are connected with the delivery of the ESFA and then to check the degree to which these skills reflect desired ESFA therapy principles, by including them in the ESFA integrity checklists, as a qualitative section, and applying the same methods as for adherence evaluation (Heilemann et al. 2014). Questionnaires and interviews could be useful methods for addressing the above. Such measures (interviews with therapy stakeholders, patient surveys and document analyses) could also be used in the future, in addition to direct observation, to make TI findings more robust (Bellg et al. 2004).

Finally, as the TI concept gains ground, a conventional criterion for the adequate level of integrity is of paramount importance; until then, decision rules can be seen as arbitrary, with inconsistency on TI score interpretation among researchers. Moreover, the

TI terminology needs to be unified for accurate interpretation of findings. Last but not least, more speech and language therapy studies need to include TI data as an essential component and for those who do so to include precise information about the methods adopted to achieve TI.

Clinical and Research Implications

This study contributes to the outcomes of the Thales aphasia project that investigated the effectiveness of ESFA aphasia therapy. The high TI scores found enhance the internal validity of the main research project, i.e. confidence that treatment outcomes relate to the treatment as originally planned, given that the protocol was implemented as planned to a high degree (Linnan and Steckler 2002). Moreover, the TI evaluation of the ESFA therapy facilitates the replication of the main study, as treatments that can be measured for adherence to protocol are likely to be sufficiently well described to be replicated (Mowbray et al. 2003, Hinckley and Douglas 2013,), and to permit future comparison across studies, both important aspects of external validity (Moncher and Prinz 1991). In addition, this study shows that ESFA as delivered in Thales is well described and therapists can effectively follow the manual and deliver the therapy as intended. Should ESFA prove to be an efficacious approach in Thales, then the first step to implementing it in clinical practice has been taken. Furthermore, the ESFA integrity checklists developed constitute the basis for a follow-up more specific TI tool that could be consistently used for future TI testing of the ESFA therapy.

Last but not least, the present study provides information about current trends in methodology for TI evaluation, while it identifies weaknesses in TI literature, especially

in the aphasia field. Overall, it contributes to the growing prominence of TI in speech and language therapy. While many researchers highlight the need for inclusion of TI data as an essential component in future speech and language therapy studies, the present study highlights the need for more emphasis on the methodological quality of TI reports, to ensure the accurate interpretation of treatment findings.

Supplemental materials

- 1. ESFA Integrity checklist for one-to-one therapy**
- 2. ESFA Integrity checklist for group therapy**
- 3. Clinicians' survey**

Table 1. Session-specific inter-rater reliability values and TI scores

Video Number	TI score R1	TI score R2	Inter- rater reliability (κ, $p < .001$)	Level of agreement (Cicchetti and Sparrow 1981)
3	87.0%	78.7%	.63	good
6	94.8%	94.4%	.94	excellent
8	96.7%	94.3%	.88	excellent
10	78.6%	75.3%	.81	excellent
14	92.6%	97.2%	.78	excellent

R1= rater 1; R2= rater 2

Table 2. Session-specific intra-rater reliability values and TI scores

Video Number	TI score T1	TI score T2	Intra- rater reliability (κ, $p < .001$)	Level of agreement (Cicchetti and Sparrow 1981)
6	94.8%	93.5%	.95	excellent
7	97.2%	97.2%	1.00	excellent
9	100.0%	97.8%	.93	excellent
11	77.2%	77.2%	1.00	excellent
14	92.6%	93.7%	.99	excellent

T1 = time 1; T2 = time 2

Table 3. Session-specific TI scores and overall TI score for one-to-one therapy approach

Session No (Therapist No)	Components Planned, Maximum TI score	Components Implemented, Actual TI score	TI score (%)
1 (1)	51	50	98%
2 (1)	61	53.5	87.7%
3 (1)	50	43.5	87%
4 (2)	34	31.5	92.6%
5 (2)	69	67	97.1%
6 (2)	67	63.5	94.8%
7 (3)	36	35	97.2%
8 (3)	61	59	96.7%
9 (3)	21	21	100%
Overall	450	424	-
Mean (SD)	50 (16.5)	47.3 (15.8)	94.6% (4.6)

Table 4. Session-specific TI scores and overall TI score for group therapy approach

Session No (Therapist No)	Components Planned, Maximum TI score	Components Implemented, Actual TI score	TI score (%)
10 (1)	98	77	78.6%
11 (1)	46	35.5	77.2%
12 (2)	55	49.5	90.0%
13 (2)	28	25	89.3%
14 (3)	81	75	92.6%
15 (3)	78	72	92.3%
Overall	386	334	-
Mean (SD)	64.3 (25.9)	55.7 (22.3)	86.7% (6.9)

Table 5. Therapists' views on complexity of ESFA

Complexity Dimensions	Therapist 1	Therapist 2	Therapist 3
Large number of (complex) behaviours required by those delivering or receiving the intervention	☒		☒
Different groups targeted by the intervention	☒		☒
There is a variability in therapy outcomes	☒		
High level of flexibility or tailoring of the intervention is permitted	☒	☒	

Figure 1. Therapists' ratings for facilitation strategies used

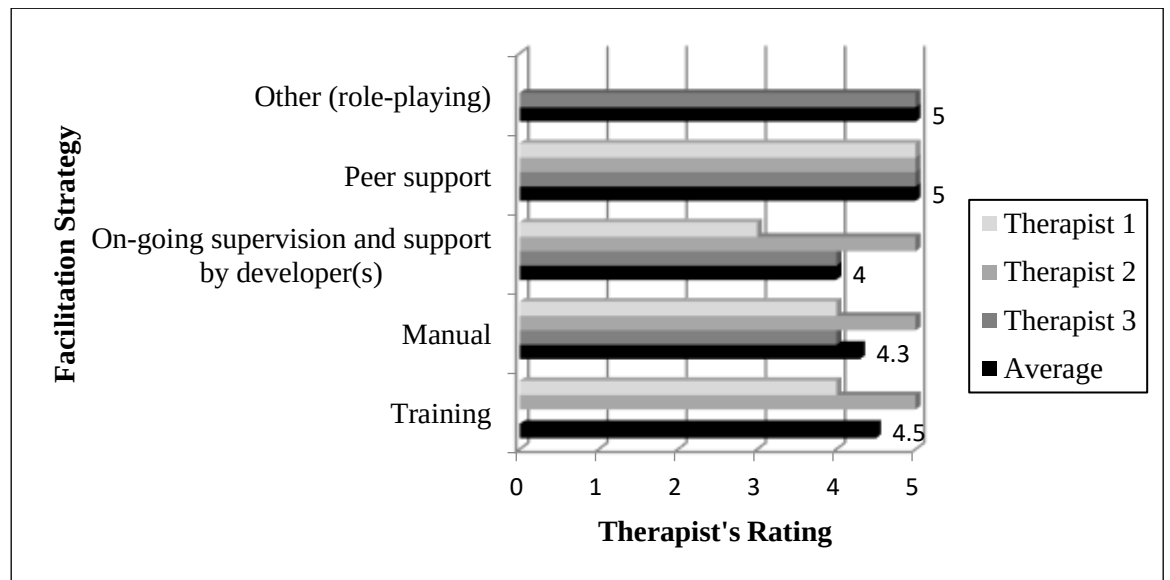
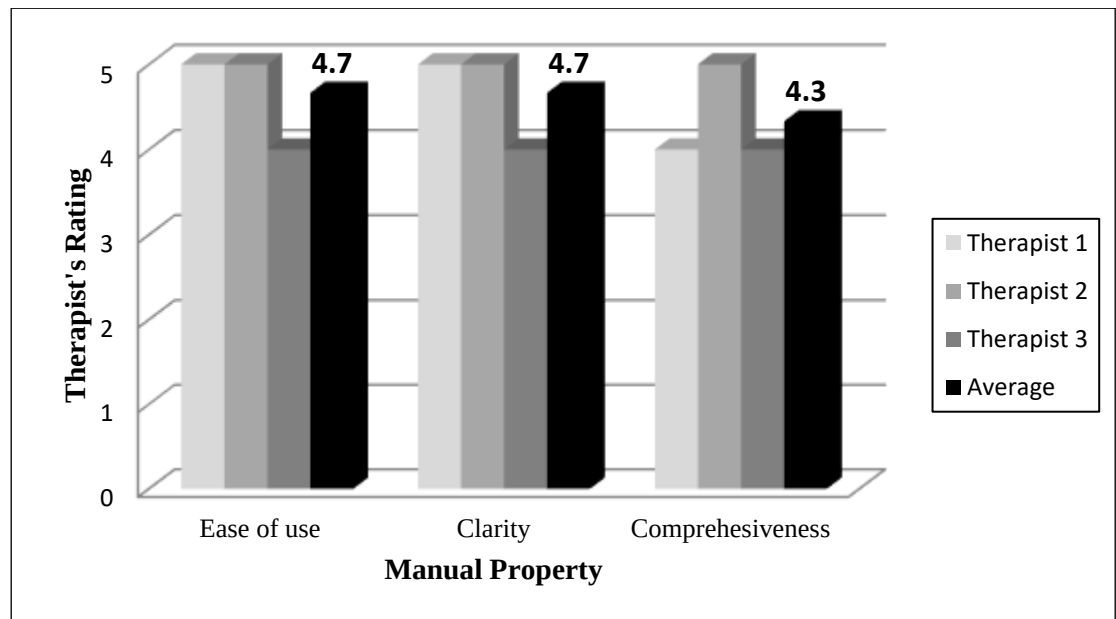


Figure 2. Therapists' ratings of the manual's properties



References

- ABRY, T., HULLEMAN, C. S., and RIMM-KAUFMAN, S. E., 2015, Using indices of fidelity to intervention core components to identify program active ingredients. *American Journal of Evaluation*, 36(3), 320-338.
- ARADI, N. S., and PIERCY, F. P., 1985, Ethical and legal guidelines related to adherence to treatment protocols in family therapy outcome research. *The American Journal of Family Therapy*, 13(3), 60-65.
- BELLG, A. J., BORRELLI, B., RESNICK, B., HECHT, J., MINICUCCI, D. S., ORY, M., OGEDEGBE, G., ORWIG, D., ERNST, D., and CZAJKOWSKI, S., 2004, Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology*, 23(5), 443-451.
- BOYLE, M., 2004, Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech-Language Pathology*, 13(3), 236-249.
- BOYLE, M., and COELHO, C. A., 1995, Application of Semantic Feature Analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology*, 4(4), 94-98.
- CARROLL, C., PATTERSON, M., WOOD, S., BOOTH, A., RICK, J., and BALAIN, S., 2007, A conceptual framework for implementation fidelity. *Implementation Science*, 2(1), 1-9.

CHERNEY, L. R., PATTERSON, J. P., RAYMER, A., FRYMARK, T., and SCHOOLING, T., 2008, Evidence-based systematic review: effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *Journal of Speech, Language, and Hearing Research*, 51(5), 1282-1299.

CHERNEY, L. R., SIMMONS-MACKIE, N., RAYMER, A., ARMSTRONG, E., and HOLLAND, A., 2013, Systematic review of communication partner training in aphasia: methodological quality. *International Journal of Speech-Language Pathology*, 15(5), 535-545.

CICCHETTI, D. V., 1994, Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290.

CLARKE, G., 1998, Intervention fidelity in the psychosocial prevention and treatment of adolescent depression. *Journal of Prevention & Intervention in the Community*, 17(2), 19-33.

COCHRANE, W. S., and LAUX, J. M., 2008, A survey investigating school psychologists' measurement of treatment integrity in school-based interventions and their beliefs about its importance. *Psychology in the Schools*, 45(6), 499-507.

COELHO, C. A., MCHUGH, R. E., and BOYLE, M., 2000, Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology*, 14(2), 133-142.

COHEN, J., 1960, A coefficient of agreement for nominal scales. *Educational and Psychosocial Measurement*, 20, 37-46.

CONLEY, A., and COELHO, C., 2003, Treatment of word retrieval impairment in chronic Broca's aphasia. *Aphasiology*, 17(3), 203-211.

CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I., and PETTICREW, M., 2008, Developing and evaluating complex interventions: the new Medical Research Council guidance. *British Medical Journal*, 337, a1655.

DIETZ, A., KNOLLMAN-PORTER, K., HUX, K., TOTH, K., and BROWN, B., 2014a, Supported reading comprehension for people with aphasia: Visual and linguistic supports. *Journal of Medical Speech-Language Pathology*, 21(4), 319-331.

DIETZ, A., WEISSLING, K., GRIFFITH, J., MCKELVEY, M., and MACKE, D., 2014b, The impact of interface design during an initial high-technology AAC experience: a collective case study of people with aphasia. *Augmentative and Alternative Communication*, 30(4), 314-328.

DUSENBURY, L., BRANNIGAN, R., FALCO, M., and HANSEN, W. B., 2003, A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research*, 18(2), 237-256.

EDMONDS, L. A., and BABB, M., 2011, Effect of verb network strengthening treatment in moderate-to-severe aphasia. *American Journal of Speech-Language Pathology*, 20(2), 131-145.

EDMONDS, L. A., and KIRAN, S., 2006, Effect of semantic naming treatment on crosslinguistic generalization in bilingual aphasia. *Journal of Speech, Language, and Hearing Research*, 49(4), 729-748.

EDMONDS, L. A., NADEAU, S. E., and KIRAN, S., 2009, Effect of Verb Network Strengthening Treatment (VNeST) on lexical retrieval of content words in sentences in persons with aphasia. *Aphasiology*, 23(3), 402-424.

EGAN, J., WORRALL, L., and OXENHAM, D., 2004, Accessible Internet training package helps people with aphasia cross the digital divide. *Aphasiology*, 18(3), 265-280.

FAROQI-SHAH, Y., FRYMARK, T., MULLEN, R., and WANG, B., 2010, Effect of treatment for bilingual individuals with aphasia: A systematic review of the evidence. *Journal of Neurolinguistics*, 23(4), 319-341.

FIXSEN, D.L., NAOOM, S.F., BLASE, K.A., FRIEDMAN, R.M., and WALLACE, F., 2005, Implementation research: A synthesis of the literature (FMHI Publication No. 231). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network. Available at: <http://ctndisseminationlibrary.org/PDF/nirnmonograph.pdf> (Accessed: 2 August 2016).

GOFF, R. A., 2013, Examining the effectiveness of intensive language action therapy in individuals with nonfluent aphasia. PhD Thesis, University of South Florida [Online]. Available at: http://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=6013&context=etd&sei-redir=1&referer=https%3A%2F%2Fscholar.google.co.uk%2Fscholar%3Fq%3DExamining%2Bthe%2Beffectiveness%2Bof%2Bintensive%2Blanguage%2Baction%2Btherapy%2Bin%2Bindividuals%2Bwith%2Bnonfluent%2Baphasia.%26hl%3Den%26as_sdt%3D0%26as_vis%3D1%26oi%3Dscholart%26sa%3DX%26ved%3D0ahU

KEwiWyOaYg6POAhWByRQKHQF_AuYQgQMIHDAA#search=%22Examining%20effectiveness%20intensive%20language%20action%20therapy%20individuals%20nonfluent%20aphasia.%22 (Accessed: 2 August 2016).

GRESHAM, F. M., MACMILLAN, D. L., BEEBE-FRANKENBERGER, M. E., and BOCIAN, K. M., 2000, Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research & Practice*, 15(4), 198-205.

GRIFFITH, J., DIETZ, A., and WEISSLING, K., 2014, Supporting narrative retells for people with aphasia using augmentative and alternative communication: Photographs or line drawings? Text or no text? *American Journal of Speech-Language Pathology*, 23(2), S213-S224.

HALLGREN, K. A., 2012, Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23-34.

HEILEMANN, C., 2013, Investigating aspects of treatment fidelity in a new conversation-based therapy for people with agrammatic aphasia and their conversation partners. PhD Thesis, University College London [Online]. Available at: https://epub.ub.uni-muenchen.de/17910/1/Master_Thesis_Heilemann_2013.pdf (Accessed: 2 August 2016)

HEILEMANN, C., BEST, W., JOHNSON, F., BECKLEY, F., EDWARDS, S., MAXIM, J., and BEEKE, S., 2014, Investigating treatment fidelity in a conversation-based aphasia therapy. *Aphasie und Verwandte Gebiete*, 2, 14-26.

HICKEY, E., BOURGEOIS, M., and OLSWANG, L., 2004, Effects of training volunteers to converse with nursing home residents with aphasia. *Aphasiology*, 18(5-7), 625-637.

HINCKLEY, J. J., and CARR, T., 2005, Comparing the outcomes of intensive and non-intensive context-based aphasia treatment. *Aphasiology*, 19(10-11), 965-974.

HINCKLEY, J. J., and DOUGLAS, N. F., 2013, Treatment fidelity: its importance and reported frequency in aphasia treatment studies. *American Journal of Speech-Language Pathology*, 22(2), S279-284.

HOFFMANN, T.C., GLASZIOU, P.P., BOUTRON, I., MILNE, R., PERERA, R., MOHER, D., ALTMAN, D.G., BARBOUR, V., MACDONALD, H., JOHNSTON, M. AND LAMB, S.E., 2014, Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*, 348,1687.

HOGUE, A., LIDDLE, H. A., SINGER, A., and LECKRONE, J., 2005, Intervention fidelity in family-based prevention counseling for adolescent problem behaviors. *Journal of Community Psychology*, 33(2), 191-211.

KADERAVEK, J. N., and JUSTICE, L. M., 2010, Fidelity: An essential component of evidence-based practice in speech-language pathology. *American Journal of Speech-Language Pathology*, 19(4), 369-379.

KAZDIN, A. E., 1986, Comparative outcome studies of psychotherapy: Methodological issues and strategies. *Journal of Consulting and Clinical Psychology*, 54(1), 95-105.

KEMPLER, D., and GORAL, M., 2011, A comparison of drill- and communication-based treatment for aphasia. *Aphasiology*, 25(11), 1327-1346.

KIRAN, S., 2008, Typicality of inanimate category exemplars in aphasia treatment: Further evidence for semantic complexity. *Journal of Speech, Language, and Hearing Research*, 51(6), 1550-1568.

KIRAN, S., and JOHNSON, L., 2008, Semantic complexity in treatment of naming deficits in aphasia: Evidence from well-defined categories. *American Journal of Speech-Language Pathology*, 17(4), 389-400.

KIRAN, S., and THOMPSON, C. K., 2003, The role of semantic complexity in treatment of naming deficitstraining semantic categories in fluent aphasia by controlling exemplar typicality. *Journal of Speech, Language, and Hearing Research*, 46(3), 608-622.

KIRAN, S., SANDBERG, C., and SEBASTIAN, R., 2011, Treatment of category generation and retrieval in aphasia: Effect of typicality of category items. *Journal of Speech, Language, and Hearing Research*, 54(4), 1101-1117

LANE, K. L., BOCIAN, K. M., MACMILLAN, D. L., and GRESHAM, F. M., 2004, Treatment integrity: An essential—but often forgotten—component of school-based interventions. *Preventing School Failure: Alternative Education for Children and Youth*, 48(3), 36-43.

LAWTON, J., JENKINS, N., DARBYSHIRE, J., HOLMAN, R., FARMER, A., and HALLOWELL, N., 2011, Challenges of maintaining research protocol fidelity in a clinical care setting: A qualitative study of the experiences and views of patients and staff participating in a randomized controlled trial. *Trials*, 12(1), 108.

LEONARD, C., ROCHON, E., and LAIRD, L., 2008, Treating naming impairments in aphasia: Findings from a phonological components analysis treatment. *Aphasiology*, 22(9), 923-947.

LICHSTEIN, K. L., RIEDEL, B. W., and GRIEVE, R., 1994, Fair tests of clinical trials: A treatment implementation model. *Advances in Behaviour Research and Therapy*, 16(1), 1-29.

LINNAN, L. and STECKLER, A., 2002, Process evaluation for public health interventions and research: An overview. In L. Linnan and A. Steckler (eds), *Process evaluation for public health interventions and research* (San Francisco: Jossey-Bass), pp. 1–23.

LONG, M. E., GRUBAUGH, A. L., ELHAI, J. D., CUSACK, K. J., KNAPP, R., and FRUEH, B. C., 2010, Therapist fidelity with an exposure-based treatment of ptsd in adults with schizophrenia or schizoaffective disorder. *Journal of Clinical Psychology*, 66(4), 383-393.

LOWELL, S., BEESON, P. M., and HOLLAND, A. L., 1995, The efficacy of a semantic cueing procedure on naming performance of adults with aphasia. *American Journal of Speech-Language Pathology*, 4(4), 109-114.

MATTIS, S., 1988, *Dementia Rating Scale: Professional manual*. (Odessa, FL: Psychological Assessment Resources).

MCINTYRE, L. L., GRESHAM, F. M., DIGENNARO, F. D., and REED, D. D., 2007, Treatment integrity of school-based interventions with children in the journal of applied behavior analysis 1991-2005. *Journal of Applied Behavior Analysis*, 40(4), 659-672.

MONCHER, F. J., and PRINZ, R. J., 1991, Treatment fidelity in outcome studies. *Clinical Psychology Review*, 11(3), 247-266.

MOWBRAY, C. T., HOLTER, M. C., TEAGUE, G. B., and BYBEE, D., 2003, Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, 24(3), 315-340.

NETEMEYER, R. G., BEARDEN, W. O., and SHARMA S., 2003, Validity. In R. G. Netemeyer (eds), *Scaling procedures: Issues and applications* (Thousand Oaks: Sage), pp. 71-87.

PAPATHANASIOU I. and MIHOU A., 2006, Elaborative Semantic Feature Analysis: A case study. [Poster]. 2006 ASHA's Convention, Miami, USA, November 16 - 18 2006.

PAPATHANASIOU I., PAPADIMITRIOU D., GAVRILOU V., and MIHOU A., 2008, Psychometric properties of BDAE in normal adult population: The effect of age and gender (Greek edition). *Psychology*, 15(4), 398-410.

PEACH, R. K., and REUTER, K. A., 2010, A discourse-based approach to semantic feature analysis for the treatment of aphasic word retrieval failures. *Aphasiology*, 24(9), 971-990.

PEREPLETCHIKOVA, F., and KAZDIN, A. E., 2005, Treatment integrity and therapeutic change: issues and research recommendations. *Clinical Psychology: Science and Practice*, 12(4), 365-383.

RAVEN, J., 2004, Coloured progressive matrices and Crichton vocabulary scale (London, England: Pearson).

RIDER, J. D., WRIGHT, H. H., MARSHALL, R. C., and PAGE, J. L., 2008, Using semantic feature analysis to improve contextual discourse in adults with aphasia. *American Journal of Speech-Language Pathology*, 17(2), 161-172.

ROSE, M., and DOUGLAS, J., 2006, A comparison of verbal and gesture treatments for a word production deficit resulting from acquired apraxia of speech. *Aphasiology*, 20(12), 1186-1209.

ROSE, M., and SUSSMILCH, G., 2008, The effects of semantic and gesture treatments on verb retrieval and verb use in aphasia. *Aphasiology*, 22(7-8), 691-706.

ROSE, M., DOUGLAS, J., and MATYAS, T., 2002, The comparative effectiveness of gesture and verbal treatments for a specific phonologic naming impairment. *Aphasiology*, 16(10-11), 1001-1030.

ROSE, M. L., RAYMER, A. M., LANYON, L. E., and ATTARD, M. C., 2013, A systematic review of gesture treatments for post-stroke aphasia. *Aphasiology*, 27(9), 1090-1127.

ROSSION, B., and POURTOIS, G., 2004, Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, 33(2), 217-236.

SCHNEIDER, S., and FRENS, R., 2005, Training four-syllable CV patterns in individuals with acquired apraxia of speech: Theoretical implications. *Aphasiology*, 19(3-5), 451-471.

SCHOENWALD, S. K., GARLAND, A. F., CHAPMAN, J. E., FRAZIER, S. L., SHEIDOW, A. J., and SOUTHAM-GEROW, M. A., 2011, Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(1), 32-43.

SNODGRASS, J. G., and VANDERWART, M., 1980, A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174-215.

STEIN, K. F., SARGENT, J. T., and RAFAELS, N., 2007, Intervention research: Establishing fidelity of the independent variable in nursing clinical trials. *Nursing Research*, 56(1), 54-62.

STUFFLEBEAM, D. L., 2000, Guidelines for developing evaluation checklists: The Checklists Development checklist (CDC). Available at: https://www.wmich.edu/sites/default/files/attachments/u350/2014/guidelines_cdc.pdf (Accessed: 2 August 2016).

SWEIFACH, J. S., and LINZER, N., 2015, Beneficence vs. Fidelity: Serving social work clients in the aftermath of catastrophic events. *Journal of Social Work Values and Ethics*, 12(1), 3-12.

WAMBAUGH, J. L., and WRIGHT, S., 2007, Improved effects of word-retrieval treatments subsequent to addition of the orthographic form. *Aphasiology*, 21(6-8), 632-642.

WRIGHT, H. H., MARSHALL, R. C., WILSON, K. B., and PAGE, J. L., 2008, Using a written cueing hierarchy to improve verbal naming in aphasia. *Aphasiology*, 22(5), 522-536.

YEATON, W. H., and SECHREST, L., 1981, Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology*, 49(2), 156-167.

YODER, P., and SYMONS, F., 2010, Interobserver Agreements and Reliability of Observational Variables. In P. Yoder and F. Symons (eds), *Observational Measurement of Behavior* (New York: Springer), pp. 159-182.

H. . Appendix H: Supplemental Material

Supplemental Material 1

Detailed Description of the ESFA therapy (One-to-one therapy approach)

Main Therapy procedures. SFA therapy as described by Boyle and Coelho (1995) was used. During the therapy session, for each item trained, the clinician initially asked the client to draw a picture from the treatment material set and then to name it. Then, presenting a semantic features chart [same as that shown in Boyle (2004), but translated in Greek language], the therapist prompted the client to think of and say words semantically related to the target word (semantic features). The chart included six categories: *superordinate category*, *use*, *action*, *physical properties*, *location*, and *association*. To elicit feature production, the therapist asked questions or provided the client with sentence-completion cues. For instance, for the *superordinate category*, a question such as “*What category does it belong to?*” was provided. Similarly, for the category *use*, a statement as “*You use it to/for _____*” was given. After the oral word production, which is the focus of ESFA therapy, the clinician prompted the client to write down the elicited features in the chart. For clients with writing difficulties, the therapist helped them to write the features with the use of an alphabet chart (e.g. pointing to the letters they needed). For clients who could not write, the therapist filled in the chart.

After the chart completion and the retrieval of the word by the client, when the SFA procedure was completed for the target word, the therapist encouraged the client

to produce phrases with the target word and each of its features. If needed, the clinician and client would say the words together or the clinician would point to the target and a feature for the client to put together in a phrase. Then, the client was encouraged to produce a sentence, including the target word and at least one of the relevant semantic features chosen by themselves. For example for the item 'table', the individual was asked during SFA to produce features such as: furniture, for dining, wooden, kitchen, chair, tea, eat, and then to elaborate these features in sentences such as: we eat at the table, we have tea at the table, the table is for dining, the table is a piece of furniture in the kitchen, etc. Elaboration of features was achieved by asking the individual to choose as many features as they wanted (one as a minimum) and to put them together into a sentence. The same strategy was followed for all treatment items. Participants had first to produce the sentence orally and then if they could to write the sentence down. It did not matter if people made errors in their sentences, e.g. syntactic or morphological errors as long as the sentence was meaningful. Help was given to participants according to their abilities; people with global aphasia for instance, needed more cues from the therapist compared to people with fluent aphasia, while with time, therapist's help was reduced. After its completion (SFA stage), the chart was used as help/ cueing as and when needed.

At the end of each session the client was asked to name all the words that had been worked on during the previous therapeutic sessions: if a target word was retrieved correctly for three consecutive sessions, without prompt or help by the therapist, and the client was able to produce correct sentences without cues or reference to the chart,

this word was removed from the therapy process and another new word replaced it. Subsequently, at the beginning of each therapy session, the client was asked to name the pictures that they had not named correctly in the previous session and to produce one sentence for each of these target words. If the client did not name the picture correctly, the chart analysis was repeated with these targets before moving on to new targets.

Additional Therapy Principles. In terms of the order of chart completion, there was flexibility. At the first therapy sessions the therapists would start for animate nouns, e.g. ‘dog’ with the first category (*superordinate* category), e.g. ‘what is it?’ or ‘what group does it belong to?’ and for inanimate nouns, e.g. ‘scissors’ with the *action* category, e.g. ‘what do we do with it?’ or the *use* category, e.g. ‘we use it for...?’, and then work their way through the other features in the following order: *physical properties*, *location*, and *association*. However, as the participants became familiar with the technique, they were let to spontaneously generate features out of sequence. When this happened, the features were written in the appropriate boxes on the chart, and if and when needed the clinician resumed eliciting features in the prescribed order, skipping over the categories that the participant had spontaneously completed. If a category was not applicable for a target word, such as when *use* and *action* categories are similar (e.g. for paintbrush: to paint), then this category was skipped by the therapist and only those deemed appropriate for the target item were elicited. If a participant named the target picture on confrontation or during the

features generation, the therapist still asked for all features to be produced, in order for the participant to build up semantic links, promoting spreading activation to related semantic concepts. This also aimed to develop feature generation as a compensatory strategy by encouraging the establishment of the technique and its use and, through repeated practice, increase the chances of a more automatic use of the technique when lexical retrieval difficulties were encountered. The client was prompted to produce as many features as possible for each category, which were then written in the category box, as the production of more related words facilitates the connections of the semantic network. Some categories encouraged more features compared to others: the *physical properties category*, for example, typically had several entries, whereas the box for *superordinate category* had fewer. The production of more than one feature for each category was not an integral component of ESFA though; one semantic feature for each category was the basic requirement. The number of the pictures worked on in each session depended on the client's performance.

During the therapy, the therapist provided cues to clients, following a specific cueing hierarchy based on the type of paraphasias produced. The hierarchy followed is demonstrated on the integrity checklists (Supplemental Materials). If the client was not able to produce the word after cueing, they were led through the entire SFA chart, with cues provided as needed, to produce the target word. When the client could not produce the target word even when all features had been listed, the clinician

produced the word orally and then the participant repeated it and named all of its features.

Supplemental Material 2

ESFA Integrity checklist (individual therapy approach): This checklist outlines the therapy process against which treatment integrity was checked. It was used by the raters for observation and rating of the individual therapy videos.

ESFA Integrity Checklist Individual therapy

Duration of session:

.....

Therapist's name:

.....

During observation of each therapy session, please evaluate whether each component was implemented (at the time point of the session that it should occur) and rate it following the instructions below. If the component is fully implemented, rate as 1; if the component is partially implemented, rate as ½; and if is not implemented, rate as 0 in the respective box. In the case that the component is not applicable, please use the initials **NA** in the respective box.

In more detail, a component should be rated as 0, when the therapist does not implement a planned component that should have been implemented, or in the case of paraphasia type, when a therapist treats a paraphasia like another one, e.g. follow the cueing hierarchy of circumlocution instead of the semantic one. A component should be rated as 0.5, when the therapist applies a planned step, but s/he does not implement it exactly as described. For example, in terms of cueing hierarchy, the therapist follows the proper hierarchy based on the paraphasia produced, but s/he adds or skips a step of the hierarchy; or s/he follows all the cueing steps, but not in the proper order (as described in the checklist). In terms of group therapy another example of a 0.5 rating could be when although a component requires therapist to involve all participants in the process, s/he applies it for only one of them. A component should be rated as 1, when the therapist implements it, exactly as described in the checklist. Finally, a component should be marked as NA, when it cannot be applied. If for example a person with aphasia produces a semantic paraphasia, the components related to phonemic paraphasias or unrelated response should automatically be marked as NA.



At the beginning of the session, when the therapist asks the client to name the pictures already analysed during the previous sessions:

1	Does the therapist offer a naming opportunity for the target word, when the card is shown?						
1i	Does the therapist control the turn taking between clients? (The one who previously produced the sentence, now produces the word)						

2	In the case that the client produces the target word with <u>phonological or phonemic paraphasias or circumlocution</u> :						
2a	Does the therapist prompt the other client (not the target one) to give cue without saying the target word?						
2b	After cueing by the other client, does the therapist offer a naming opportunity to the target client again?						
2c	If the target client still cannot produce the target word, does the therapist						

	ask the other one to name the word?						
If the other client cannot name the word:							
2d	Is phonemic cue offered by the therapist?						
2e	Is the phonemic cueing hierarchy followed, as specified in the manual? (mouthing → production of the word's first phoneme → production of the word's first syllable → production of the word's first and second syllables → production of the target word)						

2f	If therapist's cueing is not adequate for a correct production, does the chart analysis take place again?						
2g	Does the therapist ask the target client to repeat the word, after its retrieval?						
3	In the case that the client produces the target word with <u>semantic paraphasias</u> :						
3a	Does the therapist repeat the word as produced by the client and ask him/her if the word just produced is the target one?						

3b	If the client does not change his/her production, does the therapist prompt the other client to give cue without saying the target word?						
3c	After cueing by the other client, does the therapist offer a naming opportunity to the target client again?						
3d	If the target client still cannot produce the target word, does the therapist ask the other one to name the word?						

If the other client cannot name the word:							
3e	Is semantic cue offered by the therapist, if the client does not change his/her production?						
3f	Is the semantic cueing hierarchy followed, as specified in the manual? (questions semantically related with the target word → hints (This is a .../ The ...) → sentence-completion cue)						
3g	If therapist's cueing is not adequate for a correct						

	production, does the chart analysis take place again?						
3h	Does the therapist ask the target client to repeat the word, after its retrieval?						
4	In the case that the client gives <u>no/unrecognisable/unrelated response</u> :						
4a	Does the therapist prompt the other client to give cue without saying the target word?						
4b	After cueing by the other client, does the therapist offer a naming						

	opportunity to the target client again?						
4c	If the target client still cannot produce the target word, does the therapist ask the other one to name the word?						
If the second client cannot name the word:							
4d	Is cue offered by the therapist?						
4e	Is the cueing hierarchy followed, as specified in the manual? (gesture → hints (This is a .../ The ...) → sentence-completion						

	cue→mouthing→production of the word's first phoneme→production of the word's first syllable→production of the word's first and second syllables→production of the target word)						
4f	If therapist's cueing is not adequate for a correct production, does the chart analysis take place again?						
4g	Does the therapist ask the target client to repeat the word, after its retrieval?						
5	Does the therapist ask the						

	client (not the one who had to name the word) to produce a sentence with the target word?						
5a	Does the therapist prompt the other client to help him/her fill the sentence if needed?						
5b	If the target client is not able to produce a sentence after receiving help, does the therapist ask the other to produce a sentence?						
5c	If none of the clients is able to produce a sentence using the target						

	word, does the therapist help them by indicating features to be used for the sentence production?						
--	---	--	--	--	--	--	--

When a new target word is presented:

6	Does the therapist offer a naming opportunity for the target word, when the card is shown?						
6i	Does the therapist control the turn-taking between clients? (The therapist asks the clients who wants to start and if they do not reply then the therapist						

	prompts the one –not the one who produced the last sentence- to start)						
7	Does the therapist present the ESFA chart to the client?						
8	Chart completion:						
8a	Does the therapist ask the clients questions relevant to the chart categories?						
8b	Is the target client encouraged to fill in the chart by writing the indicated features under						

	the relative category?						
8c	Does the therapist prompt the other client to help the target client to fill in the chart by using the alphabet if the latter has writing difficulties?						
8d	If the client does not take initiative or struggle to start completing the chart (choosing the category s/he wants), does the therapist begin from the categories of “action” or “use”?						
8e	Are all the chart categories completed?						

	(excluding the case in which use and action receive the same reply)						
8f	If the target client is not able to give a feature, does the therapist prompt the other client to help?						
8g	Does the therapist control the turn-taking between clients?						
9	Does the therapist encourage the clients to review the chart by producing phrases including the target word and one or two features of the chart categories,						

	after its completion?						
9i	Does the therapist control the turn taking between clients?						
10	Does the therapist encourage both clients to produce a sentence, using the target word?						
10a	Does the therapist prompt the other client to help, if necessary?	Client A					
		Client B					

10b	If the client is not able to produce a sentence, does the therapist indicate and produce features that could be used and let client to complete the target word?	Client A						
		Client B						

At the end of the therapy session:

11	Does the therapist ask the clients to name all the targets that have been worked on in previous						
----	---	--	--	--	--	--	--

	therapies?						
11a	Does the therapist control the turn taking between clients? (either by having both clients name each word alternatively or by having each client name all the words in turn)						

Supplemental Material 4:. SLT participants' views on ESFA therapy

Survey

Date:

Name of respondent:

This survey aimed to investigate the views of Speech and Language Therapists (SLTs) who provided Elaborated Semantic Feature Analysis (ESFA) therapy to people with aphasia, during the Thales project, regarding different aspects of the therapy that are related to the treatment integrity (TI) concept. These views will constitute complementary to the objective data obtained through the analysis of therapy videos, and will facilitate the interpretation of findings regarding therapists' adherence to the treatment protocol and in turn the accuracy of therapy delivered.

Specifically, the present survey aims to capture therapists' thoughts in terms of a. facilitation strategies used during the therapy block to optimise the degree of treatment integrity, b. ESFA manual adequacy and usefulness for an accurate and

standardised therapy implementation and c. complexity of the ESFA therapy. To achieve this, apart from close-ended questions, open-ended questions are used to allow respondents to describe their opinion on these topics.

The following table provides definitions of key words used in this survey for a common understating of terms by all respondents.

Aspect	Definition
Treatment Integrity (including “therapists’ adherence to the treatment protocol”)	The extent to which essential components of a treatment are implemented in clinical testing as intended by treatment protocols (Dusenbury, Brannigan, Falco, & Hansen, 2003). The term refers also to the strategies that are employed to check and strengthen the consistency and precision of a therapy in order to make certain (a) that it is implemented as planned, but also (b) that over time the participants of a study receive each treatment component in a similar way (Bellg et al., 2004).
Facilitation strategies	The strategies that are used to optimize and standardize treatment integrity, <i>i.e.</i> , to ensure that everyone is receiving the same training and support, with the aim that the delivery of the intervention is as uniform as possible. Such strategies include the provision of manuals, guidelines, training, and monitoring and feedback for those delivering the intervention (Carroll et al., 2007)

Complex Intervention	Complex interventions are usually described as interventions that contain several interacting or interconnecting components, but they have several complexity dimensions that evaluators should take into account, such as: • Number and difficulty of behaviours required by those delivering or receiving the intervention • Number of groups or organizational levels targeted by the intervention • Number and variability of outcomes • Degree of flexibility or tailoring of the intervention permitted Craig et al. (2008)
---------------------------------	---

Questions

A. Facilitation Strategies

1. What facilitation strategies were used to support the accurate implementation of the therapy programme? Tick all that apply:

- ☐ Training
 - ☐ Manual
 - ☐ Ongoing supervision and support by developer(s)
 - ☐ Peer support
 - ☐ Other (please specify)
-

2. For each of the strategies you chose above, please rate their quality on a scale 1-5:

	1	2	3	4	5
Training	Very poor ☐ ☐ ☐ ☐ ☐				
	Excellent				
Manual	Very poor ☐ ☐ ☐ ☐ ☐				

Ongoing supervision and support by developer(s)	Excellent Very poor ☐ ☐ ☐ ☐ ☐
Peer support	Excellent Very poor ☐ ☐ ☐ ☐ ☐
Other (please specify)	Excellent Very poor ☐ ☐ ☐ ☐ ☐ Excellent
B. ESFA Manual	
3. a. In your opinion, does the manual describe all the essential components of the therapy process? ☐ Yes ☐ No b. If not, which therapy aspects do you think were not covered in the manual? 	

4. a. Are there any therapy components included in the manual that you think the therapist should have flexibility on how to implement, rather than follow the manual?

☐ Yes

☐ No

b. If yes, which are they?

.....

5. How would you rate the manual's properties on a scale 1-5?

	1	2	3	4	5
Ease of use					
	Very poor				
	Excellent				
Clarity					
	Very poor				
	Excellent				
Comprehensiveness					
	Very poor				
	Excellent				

6. How would you rate your adherence to the ESFA manual?

- ☐ Very low
- ☐ Low
- ☐ Moderate

☐ High ☐ Very high
C. Intervention Complexity
7. Taking into account the definition of complex interventions provided above, how would you rate the complexity of the ESFA therapy? ☐ Low (none of the complexity dimensions are applicable) ☐ Moderate (some of the complexity dimensions are applicable) Which ones? ☐ High (all of the complexity dimensions are applicable)

References

- Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., . . . Czajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol*, 23(5), 443-451. doi:10.1037/0278-6133.23.5.443
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation

fidelity. *Implement Sci*, 2, 40. doi:10.1186/1748-5908-2-40

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *Bmj*, 337, a1655. doi:10.1136/bmj.a1655

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health education research*, 18(2), 237-256. doi:10.1093/her/18.2.237

I. Appendix I: Chapter 5: Results - Descriptive Statistics Tables

5.2.3 Results on Efficacy of ESFA Therapy versus Control / Delayed Therapy

5.2.3.1. Primary Outcome Measure: Snodgrass and Vanderwart Naming Test Therapy versus Control Group

Descriptive Statistics

	Control/ Delayed Therapy Group (N=12)			Therapy Group (N=26)		
	BL1	BL2	BL3	BL1	BL2	Post
Mean	67.83	74.33	81.83	56.15	61.96	104.38
Median	79.50	88.00	96.50	47.50	57.50	105.00
Std. Deviation (SD)	57.29	62.94	69.90	45.74	49.50	73.91
Skewness	-.02	.01	.04	.33	.24	.21
Minimum	0	0	1	0	0	1
Maximum	155	164	182	150	151	238

5.2.3.2 Secondary Outcome Measures

5.2.3.2.1 Body Functions and Structure - Impairment - Based Level Results

5.2.3.2.1.1 BNT Therapy versus Control Group

	Control/ Delayed Therapy Group (N=12)			Therapy Group (N=26)		
	BL1	BL2	BL3	BL1	BL2	Post
Mean	7.75	8.92	10.00	6.85	6.81	10.50
Median	10.50	8.00	10.00	4.00	5.00	8.00
Std. Deviation (SD)	5.45	6.87	8.37	7.17	6.53	9.84
Skewness	-.42	-.02	.006	.97	1.04	1.08
Minimum	0	0	0	0	0	1
Maximum	14	17	23	25	25	34

5.2.3.2.2: Activity and Participation Level Results

5.2.3.2.2.1 ASHA-FACS Therapy versus Control Group

	Control/ Delayed Therapy Group (N=12)			Therapy Group (N=26)		
	BL1	BL2	BL3	BL1	BL2	Post
Mean	4.91	5.13	5.28	5.24	5.24	5.55
Median	4.95	5.31	5.56	5.32	5.34	5.52
Std. Deviation (SD)	1.19	1.13	1.09	1.09	1.13	.92
Skewness	-.03	.04	.08	-1.06	-.71	-.61
Minimum	2.92	3.57	3.57	1.91	2.23	3.54
Maximum	6.80	6.99	6.99	6.98	6.98	6.80

5.2.3.2.2.2 Discourse “Cookie Theft” Picture Therapy versus Control Group

	Control/ Delayed Therapy Group (N=12)			Therapy Group (N=26)		
	BL1	BL2	BL3	BL1	BL2	Post
Mean	17.65	19.16	18.64	16.35	15.22	18.14
Median	7.63	7.03	5.18	6.06	7.68	6.06
Std. Deviation (SD)	24.47	23.81	22.47	24.62	23.11	30.04
Skewness	1.39	1.09	.84	2.20	2.75	2.85
Minimum	0	0	0	0	0	0
Maximum	65.64	63.39	63.39	101.55	105.42	137.50

5.2.3.2.3 Personal Factor Level Results

5.2.3.2.3.1 GHQ-12 Therapy versus Control Group

	Control/ Delayed Therapy Group (N=12)			Therapy Group (N=26)		
	BL1	BL2	BL3	BL1	BL2	Post
Mean	6	5.50	6.17	6.27	6.04	6.12
Median	5.50	6	7	6.50	6	6
Std. Deviation (SD)	2.41	2.39	2.17	1.93	2.44	1.66
Skewness	-.28	-1.24	-.64	-.31	-.05	-.26
Minimum	1	0	2	2	2	2
Maximum	10	9	9	10	10	9

5.2.3.2.4 Quality of Life Level Results

5.2.3.2.4.1. SAQOL- 39g Therapy versus Control Group

I) Physical Domain

	Control/ Delayed Therapy Group (N=12)			Therapy Group (N=26)		
	BL1	BL2	BL3	BL1	BL2	Post
Mean	3.31	3.20	3.17	3.80	3.79	3.89
Median	3.44	3.22	3.13	4.22	4.13	4.28
Std. Deviation (SD)	1.02	1.12	.95	1.01	.98	.92
Skewness	-.31	.07	.30	-.90	-.98	-1.10
Minimum	1.56	1.56	1.94	1.44	1.38	1.31
Maximum	5	4.94	4.94	4.88	4.88	4.94

II) Psychosocial Domain

	Control/ Delayed Therapy Group (N=12)			Therapy Group (N=26)		
	BL1	BL2	BL3	BL1	BL2	Post
Mean	2.75	2.95	2.63	3.07	2.92	3.47
Median	2.69	2.72	2.59	3.16	3	3.63
Std. Deviation (SD)	.78	.77	.82	1.04	.98	.93
Skewness	.48	1.27	-.18	-.06	-.08	-.73
Minimum	1.50	2.06	1	1.31	1.31	1.31
Maximum	4.38	4.75	4.13	4.88	4.56	4.81

III) Communication Domain

	Control/ Delayed Therapy Group (N=12)			Therapy Group (N=26)		
	BL1	BL2	BL3	BL1	BL2	Post
Mean	2.52	2.83	2.65	2.77	2.82	2.86
Median	2.43	2.79	2.43	2.71	2.86	2.86
Std. Deviation (SD)	.78	1.07	1.18	.90	.90	.91
Skewness	.43	.20	.38	.32	.12	.14
Minimum	1.29	1.57	1.14	1.14	1.14	1.57
Maximum	4	4.43	4.43	4.57	4.71	4.57

IV) Overall Domain

	Control/ Delayed Therapy Group (N=12)			Therapy Group (N=26)		
	BL1	BL2	BL3	BL1	BL2	Post
Mean	2.94	3.01	2.83	3.31	3.24	3.52
Median	2.85	2.95	2.71	3.31	3.28	3.63
Std. Deviation (SD)	.60	.69	.54	.75	.73	.72
Skewness	.27	.68	.39	-.37	-.43	-1.05
Minimum	2.21	2.15	2.18	1.49	1.46	1.41
Maximum	3.97	4.46	3.77	4.67	4.49	4.56

5.2.3.2.4.2 EQ-5D Therapy versus Control Group

	Control/ Delayed Therapy Group (N=12)			Therapy Group (N=26)		
	BL1	BL2	BL3	BL1	BL2	Post
Mean	60.83	55.42	50.83	63.54	67.12	69.12
Median	60	50	60	70	70	70
Std. Deviation (SD)	23.53	20.61	15.20	19.35	16.62	15.59
Skewness	-.52	.62	-1.52	-1.26	-.23	-.24
Minimum	10	30	15	0	30	35
Maximum	100	90	65	100	100	100

5.3.3 Results on Efficacy of ESFA Direct versus Combination Approach

5.3.3.1 Primary Outcome Measure: Snodgrass and Vanderwart Naming Test

Direct versus Combination Approach

Descriptive Statistics

	Direct Approach (n=22)				Combination Approach (n=14)			
	BL1	BL2	Post	FU	BL1	BL2	Post	FU
Mean	58.91	66.23	103.64	96.32	62.14	75.29	116.79	111.64
Median	42	57.50	90.50	90.50	55.50	62	111	113.50
Std. Deviation (SD)	50.14	53.95	77.01	68.49	49.67	62.64	79.45	76.90
Skewness	.29	.11	.28	.29	.30	.41	.11	.21
Minimum	0	0	1	0	0	0	8	11
Maximum	150	151	238	218	155	182	237	224

5.3.3.2 Secondary Outcome Measures

5.3.3.2.1 Body Functions and Structure - Impairment - Based Level Results

5.3.3.2.1.1 BNT Direct versus Combination Approach

	Direct Approach (n=22)				Combination Approach (n=14)			
	BL1	BL2	Post	FU	BL1	BL2	Post	FU
Mean	6.95	7.41	10.77	10.32	7.50	8	13.14	11.21
Median	4.50	6	8	8.50	7.50	7	13.50	10
Std. Deviation (SD)	6.74	7.22	10.80	10.27	6.98	6.21	10.28	10.14
Skewness	1.02	.80	1.18	1.15	.25	.21	.20	.33
Minimum	0	0	0	0	0	0	0	0
Maximum	25	25	34	37	20	17	30	29

5.3.3.2.2 Activity and Participation Level Results

5.3.3.2.2.1 ASHA-FACS Direct versus Combination Approach

Direct Approach (n=22)					Combination Approach (n=14)			
	BL1	BL2	Post	FU	BL1	BL2	Post	FU
Mean	5.21	5.3	5.55	6.02	5.11	5.15	5.44	5.47
Median	5.22	5.27	5.52	6.10	5.58	5.58	5.42	5.84
Std. Deviation (SD)	1.12	1.08	.94	.73	1.13	1.20	.97	1.18
Skewness	-1.01	-.83	-.49	-.40	-.41	-.11	-.41	-.45
Minimum	1.91	2.23	3.63	4.56	2.92	3.61	3.54	3.50
Maximum	6.98	6.98	6.84	6.99	6.80	6.99	7	6.87

5.3.3.2.2.2 Discourse “Cookie Theft” Picture Direct versus Combination Approach

	Direct Approach (n=22)				Combination Approach (n=14)			
	BL1	BL2	Post	FU	BL1	BL2	Post	FU
Mean	16.18	14.74	17.43	17.93	17.45	16.63	18.23	17.13
Median	5.32	3.90	5.97	6.65	11.37	9.05	8.66	5.96
Std. Deviation (SD)	25.03	24.73	31.03	27.80	21.87	18.15	21.48	21.71
Skewness	2.36	2.73	3.08	2.31	1.2	.88	.94	1.06
Minimum	0	0	0	0	0	0	0	0
Maximum	101.55	105.42	137.50	109.23	63.39	48.78	60.24	58.20

5.3.3.2.3 Personal Factor Level Results

5.3.3.2.3.1 GHQ-12 Direct versus Combination Approach

	Direct Approach (n=22)				Combination Approach (n=14)			
	BL1	BL2	Post	FU	BL1	BL2	Post	FU
Mean	6.32	5.91	6	5.86	5.50	5.50	5.21	6.00
Median	6.50	5.91	6	5.86	6	6	5.50	6
Std. Deviation (SD)	1.91	2.39	1.72	1.67	1.99	2.50	2.19	1.75
Skewness	-.19	.003	-.43	-.03	-.82	-.70	-.57	-.70
Minimum	2	2	2	3	1	0	0	2
Maximum	10	10	9	9	8	10	9	8

5.3.3.2.4 Quality of Life Level Results

5.3.3.2.4.1 SAQOL- 39g Direct versus Combination Approach

I) Physical Domain

	Direct Approach (n=22)				Combination Approach (n=14)			
	BL1	BL2	Post	FU	BL1	BL2	Post	FU
Mean	3.62	3.64	3.82	3.66	3.91	3.74	3.98	3.86
Median	3.94	4.13	4.16	3.94	4	3.84	4.03	4.22
Std. Deviation (SD)	1.05	1.07	.91	1.07	.87	.98	.85	1.06
Skewness	-.90	-.90	-1.25	-.88	-.37	-.52	-.56	-.97
Minimum	1.44	1.38	1.31	1.25	2.38	2.13	2.13	1.44
Maximum	4.88	4.88	4.94	5	5	4.94	5	5

II) Psychosocial Domain

	Direct Approach (n=22)				Combination Approach (n=14)			
	BL1	BL2	Post	FU	BL1	BL2	Post	FU
Mean	2.87	2.91	3.08	3.32	3.19	3.01	3.53	3.26
Median	2.63	3	3.19	3.56	3.16	2.94	3.34	3.34
Std. Deviation (SD)	1.02	1.04	1.1	.92	.94	.73	.70	.83
Skewness	.12	.10	-.46	-.55	.10	.13	.26	-.54
Minimum	1.31	1.31	1	1.56	1.50	1.63	2.44	1.50
Maximum	4.69	4.75	4.75	4.81	4.88	4.31	4.81	4.50

III) Communication Domain

	Direct Approach (n=22)				Combination Approach (n=14)			
	BL1	BL2	Post	FU	BL1	BL2	Post	FU
Mean	2.60	2.68	2.71	2.78	2.92	3.03	3.12	3.27
Median	2.71	2.71	2.79	2.71	2.71	3	2.93	3.29
Std. Deviation (SD)	.90	.98	.97	.95	.79	.76	.94	1.15
Skewness	.53	.43	.40	.10	.29	-.28	.18	-.65
Minimum	1.14	1.14	1.43	1.14	1.71	1.71	1.57	1
Maximum	4.57	4.71	4.57	4.43	4.14	4.29	4.57	4.86

IV) Overall Domain

	Direct Approach (n=22)				Combination Approach (n=14)			
	BL1	BL2	Post	FU	BL1	BL2	Post	FU
Mean	3.13	3.15	3.31	3.35	3.43	3.29	3.62	3.49
Median	3.19	3.25	3.35	3.24	3.51	3.55	3.76	3.65
Std. Deviation (SD)	.75	.79	.78	.78	.60	.61	.57	.73
Skewness	-.07	-.08	-.46	-.33	-.07	-.37	-.55	-1.04
Minimum	1.49	1.46	1.41	1.36	2.44	2.15	2.49	1.74
Maximum	4.67	4.49	4.56	4.79	4.49	4.26	4.36	4.56

5.3.3.2.4.2 : EQ-5D Direct versus Combination Approach

	Direct Approach (n=22)				Combination Approach (n=14)			
	BL1	BL2	Post	FU	BL1	BL2	Post	FU
Mean	63.73	63.41	67.14	66.82	59.29	63.57	68.21	70.71
Median	70	62.50	70	70	60	65	70	70
Std. Deviation (SD)	18.37	18.86	17.06	13.23	25.26	19.46	17.05	14.53
Skewness	-1.26	-.13	-.19	-.52	-.48	-.004	-.32	.34
Minimum	10	30	35	40	0	30	30	50
Maximum	90	100	100	90	100	90	100	100

