CHAPTER 4

DESCRIPTION OF ACTION! APP

4.1 INTRODUCTION

In recent years, there has been much interest in the speech-language therapy and aphasia research community in the use of treatment software as part of aphasia rehabilitation. The reasoning behind this is that aphasia therapy software containing language exercises could potentially be used as an addition to regular speech and language therapy. In doing so, we can increase both the intensity and the dosage of therapy, which have been suggested to be important factors in aphasia rehabilitation (Brady, 2022; Brady et al., 2016). As is the case with regular, non-electronic speech and language therapy material, developing aphasia therapy software requires expert knowledge of both the underlying language difficulties that are present in aphasia as well as specific linguistic features that need to be addressed. It also requires making design choices for the software that are appropriate for the target population (see e.g. Messamer et al., 2016; Van de Sandt-Koenderman, 2011). In the current chapter, we describe the design of a new aphasia therapy app (Action!) and provide an overview of the app content.

4.2 ACTION!: THEORETICAL BACKGROUND AND OBJECTIVES

The cognitive processes underlying verb/sentence processing have long been an interest in language research and several theories on this topic have been proposed (see e.g. Bastiaanse & Van Zonneveld, 2004; Garrett, 1980; Levelt, 1989). These models generally assume multiple levels of processing and differ in how they propose interaction between the different levels (see De Aguiar (2015) for a discussion). In his 1989 model, for example, Levelt proposes that sentence production starts with activation of lemmas that represent concepts. These lemmas contain lexical information (word class and meaning) as well as grammatical information (e.g. for verbs, thematic roles and argument structure). For sentence processing, grammatical information related to thematic roles and argument structure is essential
for determining the appropriate number and type of arguments required for a particular verb and for integrating that verb into a sentence (see e.g. Thompson & Shapiro, 2005). For example, the number of arguments will depend on the type of verb (e.g. transitive, intransitive), while the type of arguments is essential for correct thematic role assignment (e.g. Agent, Theme) and therefore both will determine the sentence structure. After lemmas and their linked grammatical information are accessed, corresponding spoken word forms (lexemes) are activated. These spoken word forms are then inserted into a grammatical structure that is based on the lemma information and compiled by the grammatical encoder (see Bastiaanse & Van Zonneveld, 2004; Levelt, 1989). Grammatical encoding may involve movement of arguments (Chomsky, 1991, 1993), a process which results in non-canonical sentences that are more grammatically complex than their canonical counterparts (such as embedded sentences and questions; see e.g. Thompson & Shapiro, 2005). After completion of these processes, the sentence is ready to be produced by the speaker (see Figure 4.1 for a schematic overview of the Levelt (1989) model, as described in Bastiaanse & Van Zonneveld (2004)).

**Figure 4.1** Schematic overview Levelt (1989) model for sentence production, as adapted by Bastiaanse & Van Zonneveld (2004). Copyright Roelien Bastiaanse, University of Groningen
All of the processes outlined above are potentially vulnerable to breakdown in individuals with aphasia. For example, the nature of the grammatical information that is attached to verb lemmas has been argued to affect how difficult verbs are to process for people with aphasia (e.g. Bastiaanse & Jonkers, 1998; Kim & Thompson, 2000; Thompson et al., 1997). In particular, it has been found that verbs with a more complex argument structure are more difficult for people with aphasia than those with less complex argument structures (Kim & Thompson, 2000; Thompson et al., 1997). This deficit may be due to difficulty with assigning thematic roles (as reported by Barbieri et al., 2010; Webster et al., 2004), which is essential for sentence processing.

Due to its complexity, impairments to sentence processing are very common in people with aphasia (see e.g. Bastiaanse & Jonkers, 1998; Cho-Reyes & Thompson, 2012; Thompson et al., 2013). However, over the years, research has resulted in a number of well-known and effective treatment options that target verb/sentence processing impairments, such as Werkwoordsproductie op Woord-en Zinsniveau (Verb Production at the Word and Sentence Level (VWS); Bastiaanse et al., 1997), Mapping Therapy (e.g. Byng, 1988; Mitchum et al., 2000; Schwartz et al., 1994), Verb Network Strengthening Treatment (VNeST; see Edmonds, 2016) and Treatment of Underlying Forms (TUF; see Thompson & Shapiro, 2005), all of which target specific levels of processing as described in Levelt (1989). VWS consists of several treatment steps that target different stages of sentence processing, including lexical retrieval and sentence construction (though focusing only on canonical sentences; see e.g. Bastiaanse et al., 2006). Mapping Therapy targets comprehension and production through treatment of thematic role assignment (see e.g. Schwartz et al., 1994). VNeST focuses on reinforcing the links between verbs and their thematic roles, aiming to improve lexical retrieval at the lemma selection stage (Edmonds, 2016). TUF, on the other hand, focuses on the treatment of complex, non-canonical word order (as produced in the grammatical encoding stage per Levelt, 1989), which can result in generalisation of improvement to less complex sentences (see e.g. Thompson & Shapiro, 2005).

Aphasia therapy research has increasingly turned to therapy software as a complement to regular speech and language therapy (see Chapter 2 for an overview). However, despite the availability of offline verb/sentence processing treatments, the availability of digital verb/sentence processing treatment options has lagged behind other aspects of language processing (such as single word processing; see e.g. Herbert et al., 2012; Laures-Gore et al., 2021; Lavoie et al., 2019). Nevertheless, there are some pieces of therapy software that do specifically address rehabilitation of verb and sentence processing. The Dutch ACTIE! App (Bastiaanse et al., 2016), for example, provides verb and sentence processing treatment. ACTIE! was based on VWS (Bastiaanse et al., 1997), which was found to be an effective treatment in Dutch (Bastiaanse et al., 2006; Links et al., 2010), as well as in Italian (De Aguiar et al., 2015). The efficacy of the version implemented in the app has, however, not been investigated, nor is there an English version. Another example of a digital verb/sentence processing treatment is described in Furnas and Edmonds (2014), who intro-
duced a computerised version of VNeST. They found that two people with chronic aphasia improved their lexical retrieval at the word, sentence, and discourse levels after eight weeks of treatment with this software. Similarly, Thompson et al. (2010) concluded that computerised treatment based on TUF elicited treatment effects similar to that of clinician-led treatment for six people with aphasia. However, neither Furnas and Edmonds’ (2014) nor Thompson’s (2010) software was clinically available at the time we commenced software development (although a separate digital version of VNeST is available, see Tactus Therapy Solutions Ltd., n.d.). Additionally, the software described by Furnas and Edmonds (2014) and Thompson (2010) targeted one aspect of verb/sentence processing in particular, rather than more comprehensively all of the levels of processing as described by Levelt (1989). Therefore, we observed a need for English evidence-based software that targeted verb and sentence processing in a range of different contexts, from the single-word level up to complex sentences (e.g. wh-questions), capitalising on previous research from these three existing therapy programmes: VWS (Bastiaanse et al., 1997), VNeST (Edmonds, 2016) and TUF (Thompson & Shapiro, 2005).

We aimed to create a piece of software that met user needs as much as possible. We focused our attention on the needs of speech and language therapists, as clinician acceptance has been found to be an important factor for software reaching people with aphasia (see e.g. Wade et al., 2014). This decision was partially practically informed, as app development largely took place during the height of the COVID-19 pandemic, with the result that the other main user group (people with aphasia) was more difficult to reach. While there has been some previous research on clinicians’ views on aphasia therapy software (see e.g. Swales et al., 2016), there was no research looking specifically at apps. With this in mind, we conducted a survey of speech and language therapists (see Chapter 3; Cuperus et al., 2022) and were mindful of their recommendations as we developed Action!.

Finally, in line with recommendations by Van de Sandt-Koenderman (2011), our goal was to design a treatment app that could be customised by clinicians so that treatment can be maximally relevant for each individual user. We therefore aimed to create treatment steps of varying difficulty, in order to cater to varying user needs. Additionally, we wanted to include a wide range of treatment items, so that clinicians and people with aphasia could choose the precise content of treatment (both in terms of treatment difficulty and selection of relevant stimuli). Previous research (see e.g. McKelvey et al., 2010) has highlighted the importance of including functionally relevant, customisable treatment content, an aspect that was also frequently mentioned by our participants in Chapter 3 and Cuperus et al. (2022).

To summarise, our goal was to develop an aphasia therapy app focusing on verb/sentence processing whose development was guided by the following main principles: 1. The app should be evidence-based and have solid theoretical underpinnings. 2. The app should be as user-friendly as possible. 3. The app should be customisable so that treatment content is maximally relevant for individual users.
4.3 SOFTWARE DEVELOPMENT

The development of the app used the Xcode platform offered by Apple (Xcode 14, n.d.), and the Swift programming language. The therapy app was designed for use on an Apple iPad. The iOS native Core Data framework was used for in-app data storage. Core Data allows for local, offline storage of data. As all third-party iOS apps are subject to sandboxing (Apple Inc., 2021), this data can only be accessed from within the Action! app. Sandboxing therefore enhances security, as no other apps can access Action! data.

In terms of design choices, we have attempted to make the app as aphasia-friendly as possible, and, as far as possible, incorporate the recommendations of the clinicians surveyed in Cuperus et al. (2022; Chapter 3). These recommendations included using clear, modern, visually attractive animations, few visual distractions, a large and aphasia-friendly font, keeping instructions short and clear, and using a pleasant, natural voice with an appropriate (i.e. Australian) accent. Additionally, we chose to minimise the amount of typing that was required and making all on-screen text available in audio form. Consistent lay-out and formatting throughout the app were used to facilitate ease of use.

4.4 USER INTERFACE

Action! consists of two parts: 1) A clinician area for use by researchers/clinicians; and 2) a treatment area for use by people with aphasia.

4.4.1 Clinician area

A clinician using Action! first needs to set up an account by entering a user name, password, and email address. After setting up an account, clinicians are able to log on to the password protected clinician area of the app. A clinician can only access data associated with the accounts of users (people with aphasia) if the clinician’s own name was selected when the user account was set up. In this area, clinicians can customise the treatment for their clients, check their clients’ performance, administer an assessment targeted at any of the treatment steps, change their password, delete audio files, and set a daily treatment goal for their client (see overview in Figure 4.2).
4.4.2 Treatment area

People with aphasia are also required to have a user account, which can be set up with help from their therapist if needed. They are asked to enter a username, password, and select the name of their therapist from the available administration accounts registered to their device (so that their data will be shared exclusively with their own therapist). They are also asked for explicit permission to share their user data with their therapist’s account. Once their account is registered successfully, users will need to enter their user name and password to log in to Action!. This leads to a main menu in which users can choose a treatment step or view their progress (see overview in Figure 4.2).

4.5 ITEM DEVELOPMENT

Action! consists of nine different treatment steps (see more information below), which are centred around the same set of animations (with the exception of Step 6; see below). These animations depict actions and are used to treat verbs in different sentence contexts (e.g. non-finite sentences or sentences with past/future verbs).

We created a total of 151 animations using Vyond (Vyond, n.d.). The animations depict the actions described by 54 intransitive verbs, 55 non-reversible transitive verbs, and 21 reversible transitive verbs (for which we created two animations each). We chose to use
animations rather than images as this allowed for the depiction of a larger number of actions. Additionally, there is some research suggesting that dynamic depictions of verbs are easier to process by people with aphasia than static ones (e.g. Blankestijn-Wilmsen et al., 2017). We aimed to create animations that were visually as clear as possible, with limited distractions (e.g. plain backgrounds, no unnecessary props). Actors were chosen based on how they well matched the verb (e.g. a farmer digging, a burglar stealing, a child playing) as much as possible. Figure 4.3 shows a still from one of the animations that was used in the app.

Figure 4.3 Example of a still from an animation depicting the verb “steal”.

The verbs were selected from the SUBTLEX-UK database (Van Heuven et al., 2014). We sorted the verbs in the database by frequency and then worked through the 1000 most frequent verbs in order of decreasing frequency. We then selected those verbs of sufficiently high imageability to be depicted in a Vyond animation. In Vyond, there is a finite set of actions that characters can perform, which reduced the number of verbs that was included.

All animations were evaluated for acceptability by non-brain damaged adults. We initially asked ten native Australian-English speaking participants (mean age = 32.0; 7 female) to rate the first version of each animation (number of animations = 130) for how well it depicted the intended verb. In this acceptability test, verbs were presented in a sentence in non-finite form such as “the burglar is stealing money” for the example in Figure 4.3). Participants were asked to judge how well the sentence matched the animation on a scale from 0 “not at all” to 10 “perfectly”.

4
After collecting these initial ten responses, we attempted to improve the animations by making changes to those animations that were rated poorly (i.e., those with a median rating < 8.0; n = 34; keeping both versions when animations were close to the average median rating of 8.0). Some verbs were presented as several animations, so that we could select the animation rated highest for accuracy of depiction (e.g. “cleaning” had two animations: “the man is cleaning the toilet” and “the man is cleaning the windows”). Another 20 native English speakers (mean age = 31.6; 17 female) then rated the revised set of animations (n = 142). Our final selection of animations for use in the app, included only those with a median rating greater than or equal to 8.0 (average median rating = 9.51; SD = 0.64, range = 8 – 10; n = 130 rated animations; n = 54 depicting intransitive verbs, n = 55 depicting non-reversible transitive verbs, n = 21 depicting 21 reversible transitive verbs, plus an extra 21 (non-rated) animations depicting reversible transitive verbs where the two characters reversed their roles in the action). The final ratings for each animation are provided in Appendix 4.1.

### 4.6 TREATMENT STEPS

In this section, we describe the nine treatment steps available in Action!, explaining both the motivation and the design of each specific treatment step. For a briefer overview of the treatment steps, please refer to Table 5.1 in Chapter 5. However, there are some general features that apply to all steps:

- When starting a treatment step, users can either watch an explanation video (which shows an example item) or start the treatment straight away. They are not given an example item to practice themselves, but can view the explanation at any point when using the app by tapping the “explain” button that appears after tapping the “help” button (see Figure 4.4).
- The treatment steps all have the same basic screen layout, with a “help” button at the top left of the screen (which can be used to access the cueing options, and the explanation video), an instruction label in the middle, and a “tick” button at the top right of the screen (which is tapped when the user has finished with an item). The animation (in all steps except for Step 6) is in the middle of the screen. The sentence that needs to be finished is underneath the animation. After user taps “help” the “explain” and “cues” options appear at the bottom of the screen (see Figure 4.4).
- Each step contains a range of cueing options that can be used as needed. The cueing buttons appear at the bottom of the screen after the user taps “help” (see Figure 4.4), allowing the user to select any cue. When a cue type is tapped, the sentence to be completed label underneath the animation changes to provide details of that particular cue. This cue will remain visible until the user chooses a different cue or taps the tick button. Cues vary in nature per step, with semantic, phonological, and orthographic
Figure 4.4 Basic layout of treatment screen (for item “driving”)

Note: The bottom row of (explain/cue) buttons only becomes visible when requested via the “help” button. Speaker icons can be tapped to hear audio of the accompanying text.
and sentence cues available (see details per step below; Boyle, 2004; Lorenz & Nickels, 2007; Meteyard & Bose, 2018; Stimley & Noll, 1991). Cues can be used an unlimited number of times for the same item.

- Users can re-watch any animation by tapping it.
- Any text in the app can be read out by tapping the speaker icon next to the text (see e.g. Figure 4.4).
- Missing words in sentences (e.g. “the man is _____ his teeth”) are pronounced as “hmm-hmm” (e.g. “the man is “hmm-hmm” his teeth”). This sound was consistently two syllables, regardless of the number of syllables of the missing word. This is not the case for cue sentences that provide the first verb sound (e.g. “the man is b_____ his teeth”, in which case “b_____” is pronounced as /b/.
- Previous research has found that self-monitoring can be relatively intact in people with agrammatic aphasia (see e.g. Schuchard et al., 2017; Schwartz et al., 2016), consequently, in steps 1-4, to allow clinicians to monitor performance, users are asked to judge their accuracy. After completing an item, users are taken to a separate screen which shows the correct response for the completed item. They are then asked to judge whether they did or did not give the correct answer (by tapping a “yes” or “no” button; see Figure 4.5). A “win” sound is played when they tap “yes” (an instrumental sound similar to positive beeps used in videogames). No negative feedback is provided when tapping “no”. After providing a “yes” or “no” response, the app automatically moves on to the next item (again presented in the format as per Figure 4.4).
- In Steps 5-9, when the user provides a response that matches the target, they hear a “win” sound.
- Items are presented in random order. For steps 1-4, users can move on to the next item without supplying an answer should they wish to, simply by tapping the tick button without giving an answer. Presentation of these skipped items is recorded, so that SLTs are aware of how often this choice is made. In the remaining steps, users can only continue to the next item once they have provided an answer.
- Users can stop halfway through a treatment step. If they do not terminate the app entirely (but leave it running in the background, i.e. by tapping the home button) they will continue where they left off at a later point. If users do terminate the app, it will start at the log-in screen on reopening (in which case they will restart when selecting a treatment step, with earlier progress being lost).
- At the end of a treatment step, users are given a feedback graph that shows how much time they have spent on treatment, as well as their daily treatment time target, which can be set by their clinician.
4.6.1 Step 1 – Naming actions with verbs in isolation

In this step, the user sees an animation and is asked to name the action using a verb in isolation (present continuous, -ing, form; see Figure 4.6). This step includes intransitive and non-reversible transitive verbs and requires a single-word response.

Motivation

Action naming is frequently found to be impaired in people with aphasia (see e.g. Bastiaanse & Jonkers, 1998; Marshall et al., 1998). Verb retrieval is often more severely impaired compared to nouns (due to extra complexity associated with verbs including the grammatical information attached to verb lemmas, see e.g. Barbieri et al., 2010; Kim & Thompson, 2000). These difficulties may also point to a more general word retrieval problem, which may indicate an impairment in (access to) lemmas or lexemes as represented by Levelt (1989). Treatment step 1 requires people to use conceptual information from the animation in order to retrieve lexical forms, thereby targeting the first step in the Levelt (1989) model. Step 1 also serves as a preparation for the next steps, allowing the user to get used to the Action! environment. It is based on Treatment Step 1 in the Dutch ACTIE! Protocol (Bastiaanse et al., 2016). This type of single-word verb retrieval treatment has been found to be effective for improving verb retrieval in people with aphasia (see e.g. Webster & Whitworth, 2012).
Figure 4.6 Screenshot of Step 1 item “crying”

Description
1. The animation is shown on the screen. The user is asked to describe “what is happening?”.
2. The user can say the verb aloud or choose any cues they wish by tapping “help” and selecting the desired cue. Cues that are given in this step are:
   a. Cloze sentence: Replacing the text on the “sentence to be completed” label underneath the animation in Figure 4.4, the user hears and sees a high probability cloze sentence (which is different to the target sentence) and which can be completed using the present continuous (-ing) form of the verb (e.g. “the boy must be very sad. I saw him ______” for “crying”).
   b. First sound/letter: The user hears the first phoneme of the verb and sees the first letter on the screen (e.g. “c______” for “crying”)
   c. Repetition/reading: The user hears the correct -ing form of the verb and sees the verb on screen. (e.g., “crying” for “crying”)
3. The user says the verb aloud and taps the tick button (indicating they have produced a response).
4. The user is taken to the next screen to judge their performance. They are presented with the correct verb in auditory and written form (text + audio) and asked to judge whether this is identical to the answer that they provided (see Figure 4.5 above).
5. The user then moves on to the next item.
4.6.2 Step 2 – Inserting present continuous forms in sentences

In Step 2, the user is required to view the animation, while listening to and seeing a sentence that is missing a verb (e.g. “the man is _______ his teeth”). They are asked to complete the sentence using the present continuous form of the relevant verb (e.g. “brushing”; see Figure 4.7). This step includes intransitive and non-reversible transitive sentences.

Motivation

Some authors argue that since verbs are used to build a sentence, it makes more sense to treat them in a sentence context rather than in isolation (e.g. Bastiaanse et al., 2006; Links et al., 2010). This step is therefore similar to Step 2 of the Dutch VWS/Dutch ACTIE! (Bastiaanse et al., 1997, 2016), in that it requires sentence completion and thereby targets the syntactic level. It requires mapping of thematic roles onto grammatical roles, using the retrieved lemma information as per Levelt et al. (1989; see also Bastiaanse et al., 2006) and therefore also targets grammatical encoding. However, in contrast to the Dutch ACTIE! and its Italian counterpart (De Aguiar et al., 2015), we do not elicit an infinitive, but rather the present continuous (-ing) form. We made this choice because the present continuous is the verb form that is usually elicited and trained in action naming in English and is relatively easy for people with verb and sentence processing impairments to produce (see e.g. Williams & Canter, 1987). Treating verb retrieval in sentence context has been found to induce retrieval improvements at the single word and sentence level (Webster & Whitworth, 2012).

Figure 4.7 Screenshot of Step 2 item “brushing”
Description

1. The animation is shown on the screen. The user is instructed to “repeat the sentence and finish it correctly”. The sentence to be completed is presented in auditory form and written at the bottom of the screen. The sentence should be completed with an -ing form (e.g. “the boy is ______” for “the boy is smiling” or “the man is ______ his teeth” for “the man is brushing his teeth”)

2. The user can say the sentence aloud or choose any cues they wish by tapping “help” and selecting the desired cue. Cues that are given in this step are:
   a. Cloze sentence: The user hears and sees a high probability cloze sentence which is different to the target sentence and which can be completed using the present continuous form of the verb. (e.g. “my hair is full of knots. It needs ______” for “brushing”)
   b. First sound/letter: The user hears the first phoneme of the verb and sees the first letter on the screen (e.g. “the man is b______ his teeth” for “the man is brushing his teeth”.
   c. Repetition/reading: The user is provided with the auditory and written form of the full sentence to repeat and/or read (e.g. “the man is brushing his teeth”).

3. The user says the completed sentence (with or without cues) and taps the tick button (indicating that they have produced a response).

4. The user is taken to the screen to judge their performance. They are presented with the correct sentence in auditory and written form and asked to judge whether this is identical to the answer they provided (see Figure 4.5 above).

5. The user then moves on to the next item.

4.6.3 Step 3 – Sentences with verbs in habitual present tense

The user is required to repeat an incomplete sentence, that starts with “every day”, and complete it using the habitual present (-s) form of the verb (see Figure 4.8).

Motivation

Previous research has found that using finite verb forms (especially those referring to the past) can be difficult for people with agrammatic aphasia (Bastiaanse et al., 2011). This step is an introduction of the use of finite verb forms and was adapted from Step 3 in the Dutch ACTIE! Treatment protocol (Bastiaanse et al., 2016). It should be noted that this step has an added complexity in Dutch. This is due to a phenomenon in which finite verb forms are moved from the end of the clause to the second position in Dutch, known as V2-movement (see Zwart, 2011). Since there is no V2-movement in English, in Action! this treatment step therefore is easier than in Dutch and functions as an introduction into the use of finite verb forms, similar to Step 3 in the Italian adaptation of ACTIE! (see De Aguiar et al., 2015). As such, this step requires verb inflection, using morphosyntactic information that is activated when lemmas are retrieved from the lexicon (see Levelt 1989) and the retrieval
of the corresponding phonological representations, as well as grammatical encoding. Treating verbs in this context is particularly appropriate as previous research has found that targeting tense morphology in the context of verb retrieval enhances the chances of generalisation to untreated verbs occurring (De Aguiar et al., 2016).

**Figure 4.8** Screenshot of Step 3 item “smiles”

![Screenshot of Step 3 item “smiles”](image)

**Description**

1. The animation is shown on the screen. The user is asked to “repeat the sentence and finish it correctly”, hearing and seeing a sentence with the verb missing (e.g. “every day, the boy _____” for “every day, the boy smiles” or “every day, the woman _____ popcorn” for “every day, the woman sells popcorn”). The sentence needs to be completed with the habitual present verb form (-s).

2. The user can say the sentence aloud or choose any cues they wish by tapping “help” and then selecting the desired cue. Cues that are given in this step are:
   a. Cloze sentence: The user hears and sees a high probability Cloze sentence which can be completed using the present continuous form of the verb (same as in Step 1 and 2). Note, this cue assists lexical retrieval, however, as the form prompted is present continuous, and this step requires the habitual present the user will still need to conjugate the verb correctly in the context of the target sentence.
   b. First sound/letter: The user hears and sees the sentence with the first phoneme/
letter of the missing verb added (e.g. “every day, the boy s_____” for “every day, the boy smiles”.

c. Repetition/reading: The user hears and sees the entire sentence (e.g. “every day, the boy smiles”.

3. The user says the completed sentence (with or without cues) and taps the tick button (indicating that they have produced a response).

4. The user is taken to the next screen to judge their performance. They are presented with the correct sentence in spoken and written form and asked to judge whether this is identical to the answer that they provided (see Figure 4.5 above).

5. The user then moves on to the next item.

**4.6.4 Step 4 – Sentences with verbs in past and future tense**

In this treatment step, the user is asked to complete sentences with a verb appropriate for the time referred to: in the past (“yesterday”) or in the future (“tomorrow”; see Figure 4.9).

**Motivation**

Making references to past and future events is important for everyday communication. The PAst DIscourse LIinking Hypotheses (Bastiaanse et al., 2011) states that referencing to the past is particularly difficult as it requires discourse linking (i.e. a relationship between speech and a past event needs to be established). Indeed, past time reference has been found to be impaired in people with aphasia (e.g. Abuom & Bastiaanse, 2013; Jonkers & de Bruin, 2009). People with aphasia may also, however, struggle with future time reference (e.g. Nanousi et al., 2006; Rofes et al., 2014; Siriboonpipattana et al., 2020). The two tenses are treated simultaneously in this treatment step in accordance with Step 4 of the ACTIE! Treatment protocol (Bastiaanse et al., 2016), which in turn was adapted from Step 3 of the Italian version of ACTIE! (De Aguiar et al., 2015). As the sentence frame is provided to users, the focus in this step is on retrieval of the correct morphemes (tense inflections) and combining these correctly with the retrieved lexical forms, as well as inserting the verb in the sentence frame (thereby targeting lemma and lexeme retrieval, as well as grammatical encoding; see Levelt, 1989). As with Step 3, treating verb retrieval as well as tense morphology may increase chances of increased lexical retrieval of untreated verbs (see e.g. De Aguiar et al., 2016).

**Description**

1. The animation is shown on the screen. The user is asked to “repeat the sentence and finish it correctly”, hearing and seeing a sentence with the verb missing (e.g. “yesterday, the man _____ the toilet” for “yesterday, the man cleaned the toilet” or “tomorrow, the girl _____” for “tomorrow, the girl will jump”). The sentence needs to be completed with the past or future tense verb form, depending on the sentence.

2. The user can say the sentence aloud or choose any cues they wish by tapping “help”
and selecting the desired cue. Cues that are given in this step are similar to those in the Dutch and Italian versions of ACTIE! (De Aguiar et al., 2015):

a. Verb: This cue gives the user the verb stem (e.g. “clean”) in written and auditory form.

b. Multiple choice: This cue provides the user with the written and auditory forms of the missing verb in past and future tense (e.g. “cleaned”, “will clean”), which appear in random order on two buttons, on the screen underneath the sentence that needs to be completed. The user is asked to tap the correct verb form. When they select the correct verb form (by tapping the button), the button with the correct verb turns green and they hear a “win” sound. When they select the incorrect verb form, the button with the incorrect word form turns red while they hear a “loss” sound. The user is auditorily encouraged “try again”.

3. The user says the sentence out loud and taps the tick button (indicating that they have produced a response).

4. The user is taken to the next screen to judge their performance. They are presented with the correct sentence in auditory and written form and asked to judge whether this is identical to the answer that they provided (see Figure 4.5 above).

5. The app then moves on to the next item.
4.6.5 Step 5 – Thematic roles - non-reversible sentences

This step focuses on non-reversible transitive sentences. The user is asked to order the different constituents of the sentence (see Figure 4.10). The aim of this treatment step is to create awareness of the different sentence constituents and their thematic roles.

Motivation

People with agrammatic aphasia tend to struggle with thematic role assignment, which is crucial for comprehension and production of sentences (see e.g. Rochon et al., 2005; Schwartz et al., 1994). This observation gave rise to the development of Mapping Therapy, a type of treatment which targets the assignment of thematic roles in the processes of sentence comprehension and production (see e.g. Rochon et al., 2005). Improvements in sentence comprehension as well as production have been previously reported to occur as a result of Mapping Therapy (see e.g. Mitchum et al., 2000). Action! treatment Step 5 follows this research tradition by initially focusing on the construction of non-reversible sentences. While the sentence constituents are provided, the user is required to subsequently retrieve the necessary lexical and grammatical information required to build the argument structure. They are also required to access the correct phonological representations and insert these in the correct position in the sentence frame. This step therefore targets lemma and lexeme retrieval, as well as grammatical and phonological encoding as described in the Levelt (1989) model. A similar treatment step (slightly different as it was combined with Step 4) was included in the Italian version of ACTIE! (De Aguiar et al., 2015). The Italian ACTIE! treatment was found to improve lexical retrieval for treated as well as untreated verbs, which is in line with the positive effects of mapping therapy as reported in previous literature (see e.g. Mitchum et al., 2000).

Description

1. The animation is shown on the screen. Below the sentence are three blank lines representing the verb and the noun phrases around it (e.g. “the chef is cutting fish” would be: “_____  _____  _____”) At the bottom of the screen are the word(s) that are required to fill in the blanks (e.g. “the chef”, “is cutting”, “fish”). These three constituents are presented in random order. The user is instructed to drag the words at the bottom to the line representing the correct sentence position.

2. The user can use any cues that they may need (the cue buttons appear underneath the sentence words i.e. “is cutting”, “fish”, and “the chef” in Figure 4.10). Cues that are given in this step are:
   a. “Who cue”: provision of position and identity of the agent – both the word and the line light up and change to the same colour and remain in this colour until the user moves to the next item.
b. “Action cue”: provision of position and identity of the verb – both the word and the line light up and change to the same colour and remain in this colour until the user moves to the next item.

c. “What cue”: provision of position and identity of the theme – both the word and the line light up and change to the same colour and remain in this colour until the user moves to the next item.

3. The user drags each sentence segment to the blank line. If correct, the segment will stay in this position. If incorrect, the sentence segment moves back to its original position. This continues until the user has dragged all sentence segments to their correct positions (using cues if needed).

4. Once all sentence segments are correctly positioned, the complete sentence (which is now visible where the blank lines were presented) is presented in auditory form. The user is asked to repeat/read the sentence.

5. The user taps the tick and this leads to the presentation of the next item.
4.6.4 Step 6: Thematic roles – expanding verb concepts

In this step, we aim to strengthen the user’s awareness of thematic roles surrounding the verbs. The user is asked to consider a variety of phrases representing thematic roles, and to put them in their correct position relative to the verb (see Figure 4.11).

Motivation

This step is based on VNeST (Edmonds, 2016). VNeST aims to focus explicitly on the central role that the verb has in a sentence. Edmonds (2016) proposes that by using the verb with a range of different agents and themes, VNeST emphasises the range of semantic concepts that can be used around a verb. In doing so, it aims to promote generalisation and understanding of thematic roles. As the central position and lexical information of the verb is provided, the focus in this step is on assigning phonological representations to the correct position in the sentence frame based on their thematic roles (as per Levelt, 1989). Such therapy has previously been found efficacious in offline form (see Edmonds, 2016, for a review). A previous digital version of VNeST similarly found word retrieval gains at the word, sentence, and discourse level after eight weeks of treatment (Furnas & Edmonds, 2014).

Figure 4.11 Screenshot of Step 6 item “painting”
This step is different from the others in that the user does not see an animation. It is also unique in that there are multiple correct answers. Any combination of agent + verb + theme is considered correct, regardless of how realistic the sentence would be (e.g. “the artist is painting the wall” is considered correct even if “the painter is painting the wall” might be more likely). Our reasoning is that it is hard to definitively argue that a sentence is incorrect simply because we consider it is less likely than a different combination (for example, an artist could be employed to paint a mural on a wall). However, “who” items are only acceptable in the “who” column and “what” items are only accepted in the “what” column (e.g. “the wall is painting the artist” is not acceptable).

For our treatment protocol, this step is relevant for two reasons. Firstly, it emphasises that verbs can be used with different arguments, which is a key skill for everyday conversation. Secondly, promoting understanding of thematic roles and their relation to the verb is essential for the next steps, which focus on movement of arguments.

**Description**

1. A verb is presented in the middle of the screen. On the left-hand side is a “who” column and on the right-hand side is a “what” column. At the bottom of the screen are several words that the user needs to drag to the most appropriate position.

2. The user can use any cues they may need:
   a. “Who cue” – This cue shows the user which of the words at the bottom would usually be considered to go in the “who” column. These words and their respective positions light up in the same colour.
   b. “What cue” – This cue shows the user which of the words at the bottom would usually be considered to go in the “what” column. These words and their respective positions light up in the same colour.

3. The user drags the word to a blank line in the who or what column. If correct, the word will stay in this position. If incorrect, the word moves back to its original position.

4. When all words are matched correctly, the user is instructed to read the sentences aloud and to tap the tick button when they have finished.

5. The app then moves on to the next item.

**4.6.7 Step 7 – Thematic roles - reversible sentences**

This step is identical to Step 5 in that it focuses on the thematic roles of the verb in the process of sentence construction, except it presents reversible instead of non-reversible sentences (see Figure 4.12).

**Motivation**

Reversible sentences are particularly difficult to process for people with aphasia because of the difficulties they have with thematic role assignment (see e.g. Friedmann & Shapiro, 2003; Thompson & Shapiro, 2005. Whereas in Step 5 people can rely on semantic knowledge...
to determine thematic role assignment (e.g. a chef can cut fish but a fish cannot cut a chef), Step 7 requires the user to rely solely on conceptual input (the animation) in order to derive thematic roles accurately and assign them to the correct position in the sentence frame (e.g. a chef can push a nurse but a nurse can also push a chef, meaning that semantic knowledge alone is not sufficient to create a sentence). Previous research has found that treating reversible sentences in a Mapping Therapy-based context can lead to sentence processing improvements, with the potential for gains to generalise to untrained structures as well as narrative production (Rochon et al., 2005).

Figure 4.12 Screenshot of Step 7 item “pushing”

Description
1. An animation is shown on the screen. Below the animation are three blank lines representing the verb and the noun phrases around it (e.g. “the nurse is pushing the chef” would be: “_____ _____ _____”). At the bottom of the screen are the phrases that are needed to fill in the blanks (e.g. “the nurse”, “is pushing”, “the chef”). These three constituents are presented in random order. The user is instructed to drag the words at the bottom to the line representing the correct sentence position.
2. The user can use any cues they wish. Cues that are given in this step are:
   a. “Who cue”: provision of position and identity of the agent – the word and the line change to the same colour and remain in this colour until the user moves to the next item.
b. “Action cue”: provision of position and identity of the verb – the word and the line change to the same colour and remain in this colour until the user moves to the next item.

c. “To whom cue”: provision of position and identity of the theme – the word and the line change to the same colour and remain in this colour until the user moves to the next item.

3. The user drags the sentence segment to one of the blank lines. If correct, the segment stays in this position. If incorrect, the sentence segment moves back to its original position. This continues until the user has dragged all sentence segments to their correct positions (using cues if needed).

4. Once all sentence segments are correctly positioned, the complete sentence (which is now visible where the blank lines were presented) is automatically presented in auditory form. The user is asked to repeat/read the sentence.

5. The user taps the tick and the app continues to the next item.

4.6.8 Step 8: Who questions

The user is taken through the steps required to create a “who-question” from a declarative sentence (see Figure 4.13).

Motivation

The ability to ask questions is often impaired in people with aphasia (e.g., Caramazza & Zurif, 1976). The degree of impairment is influenced by the complexity of the question. Since subject wh-questions follow the canonical word order of English, these questions are relatively simple to produce (Shapiro, 1997) and require grammatical as well as phonological encoding (as described in Levelt et al. 1989). This treatment step was partly based on the Treatment of Underlying Forms (TUF) therapy programme (Thompson & Shapiro, 2005). Similar to TUF, we use explicit explanations of the linguistic processes that underlie question formation.

Description

1. The agent and theme of the sentence (e.g. “the chef” “the nurse”) are shown underneath the animation, with auditory versions available through tapping the speaker symbols next to the words (see Figure 4.13). The user is asked who is doing the action and to tap on the corresponding word. If the user selects the wrong word, they hear a negative sound and are encouraged through auditory instruction to tap the other word.

2. When the user selects the subject correctly, a sentence with a blank line at the end appears (e.g. the nurse is painting the chef _______ underneath the animation (see Figure 4.14).
Simultaneously, two labels with “who” and “?” appear at the bottom of the screen (see Figure 4.14; these can be tapped to hear an auditory version). The user is auditorily instructed to drag these labels to the correct sentence position, either on the blank line (for “?”), or over existing text (for “who”). When they do, the existing text underneath the animation (“the nurse” and “_____” in Figure 4.14) disappears, and is replaced with “who” and “?” respectively. This continues until both labels have been dragged to the correct position and the who question (e.g. “who is painting the chef?”) is visible in its entirety underneath the animation. If the user drags either “who” or “?” label to an incorrect position, the label “who” or “?” is automatically returned to their initial position (as shown in Figure 4.14). Users are also provided with explicit auditory feedback in the case of any errors (e.g. “the question mark goes at the end of the sentence”).

The user can use any cues they wish (these are only accessible after identification of the subject):

a. “Who cue”: provision of the intended position of “who” in the sentence – who and the text referring to its intended referent change to the same colour and remain in this colour until the user moves to the next item.
b. Get positions – provision of the identity and intended location of both “who” and the question mark; both words and their respective locations change to the corresponding colours and remain in this colour until the user moves to the next item.

5. The user is asked to repeat the question and tap the tick when they have finished.
6. The app then moves on to the next item.

**Figure 4.14** Screenshot of Step 8 item “painting”– dragging subject and question mark to their correct positions

4.6.9 Step 9: What questions

The user is taken through the steps required to create a “what-question” from a declarative sentence (see Figure 4.15).

**Motivation**

Object (“what”) wh-questions are more complex than subject wh-questions as they do not adhere to the canonical (SVO) word order, as a result of syntactic movement (Thompson, 2016). In terms of the Levelt (1989) model, this means that this treatment step requires the same processes as Step 8 (grammatical and phonological encoding), with syntactic movement and non-canonical word order adding complexity. Similar to the previous step,
we aimed to incorporate the principles of the TUF protocol (Thompson & Shapiro, 2005). The user is given explicit feedback on how to formulate a what-question and is encouraged to verbally produce the question. Previous studies in the TUF framework have found that this approach can lead to substantial improvements in treated what-questions, with potential generalisation to untrained what-questions and who-questions (see Thompson & Shapiro, 2005 for an overview).

**Figure 4.15** Screenshot of Step 9 item "stealing" – picking the “what” word

**Description**

1. The agent and theme of the sentence (e.g. “the burglar” and “money”) are displayed underneath the animation (see Figure 4.15), with auditory versions available by tapping the speaker symbols next to the words. The user is asked to tap the “what” word (theme). If the user selects the wrong word, they hear a negative sound and are encouraged to tap the other word.

2. When the user selects the object correctly, a sentence with three blank lines will appear (e.g. _____ _____ the burglar stealing _____”) underneath the animation (see Figure 4.16).
3. Simultaneously, three labels with “what”, “is” and “?” appear at the bottom of the screen (see Figure 4.16; these can be tapped to hear an auditory version). The user is instructed (auditorily and in writing) to drag these labels to the correct sentence position (i.e., onto a blank line). When they do, the existing blank line (“_______”) disappears, and is replaced with “what”, “is” or “?” respectively. This continues until all labels have been dragged to the correct position and the what-question (e.g. “what is the burglar stealing?”) is visible in its entirety underneath the animation. If the user drags the labels to the incorrect position, they are automatically moved back to their initial position (as shown in Figure 4.16). Users are also provided with explicit feedback in the case of any errors (e.g. “the question mark goes at the end of the sentence”).

4. The user can use any cues they may need:
   a. “What cue”: provision of position of “what” – what and its intended location change to the same colour and remain in this colour until the user moves to the next item.
   b. Get positions – provision of the identity and location of “what”, “is”, and the question mark; these words and their respective locations change to corresponding colours and remain in this colour until the user moves to the next item.

5. The user is asked to repeat the question and tap the tick when they have finished.

6. The app then moves on to the next item.

**Figure 4.16** Screenshot of Step 9 item “stealing” – dragging the labels to their correct positions
This chapter introduced the Action! app and described its rationale and development. Action! allows people with aphasia to independently work on verb and sentence processing, while supervised by their therapist. As a result of the different treatment steps that the app offers, individuals with aphasia can work at a level that is tailored to their needs. This app will meet clinical needs of people with aphasia and speech and language therapists alike.

As described above, the development of Action! was guided by three main principles. Firstly, we aimed to create an app that was evidence-based. We have done so through incorporating elements from different (offline) treatment programmes, such as VWS/ACTIE (see e.g. Bastiaanse et al., 1997, 2016), mapping therapy (see e.g. Byng, 1988; Mitchum et al., 2000; Schwartz et al., 1994), VNeST (see e.g. Edmonds, 2016; Furnas & Edmonds, 2014), and TUF (see e.g. Thompson et al., 2010; Thompson & Shapiro, 2005). These well-researched treatment programmes form the backbone of our app and have allowed us to create a therapy app that targets many of the different processes described by Levelt’s model for sentence production (1989). Secondly, our goal was to make Action! as functionally relevant as possible. Through incorporation of treatment steps of varying difficulty, as well as customisable treatment content (as recommended by, for example, Van de Sandt-Koenderman (2011), Action! can be personalised for targeted treatment and includes highly frequent verbs that were selected using SUBTLEX-UK (Van Heuven et al., 2014). Finally, we have attempted to make Action! as user-friendly as possible, by reference to the existing literature (Swales et al., 2016) and by conducting a survey among one of our main user groups (i.e. speech and language therapists; see Cuperus et al. 2022; Chapter 3). This survey resulted in useful feedback that we could implement in the app, such as using short and clear instructions recorded by a native speaker with a location-appropriate accent (i.e. Australian English).

Action! in its current form remains in a relatively early stage of development. In the future, we could consider increasing the number of treatment steps (potentially incorporating different types of questions and/or complex sentences that are not currently addressed) as well as increasing the verbs and animations available as treatment items. Importantly, piloting of Action! With our intended users (i.e. people with aphasia and clinicians) would be critical to provide us with feedback. Such feedback will facilitate refinement of treatment steps and design principles more broadly. Piloting will also allow us to understand Action!’s potential to provide effective treatment (see Chapter 5 for some initial results). Nevertheless, Action! Is a thoroughly planned and well-developed product based on solid theoretical background and existing therapy programmes. It incorporates features that are not possible with pen-and-paper exercises (such as the use of animations and having all user data (including audio recordings) available to clinicians) and can be used independently by people with aphasia. While we hope that Action! Can be further developed and researched in the future, we believe that, in its current form, it may already be a useful addition to regular speech and language therapy.