

*Anna Konieczna*

*PWSZ Nysa*

---

## **RAPORT WERBALNY I JEGO WYKORZYSTANIE W BADA- NIACH NAD STRATEGIAMI ROZWIĄZYWANIA TESTÓW Z ZAKRESU CZYTANIA**

### **The verbal report and its use in research on test-taking strategies in reading**

The article discusses kinds of verbal reports and their characteristic features. More specifically, it focuses on: a) the validity of particular types of verbal reports and the data these are likely to provide, b) a quite difficult distinction between introspection and retrospection in researching reading, and c) the characteristic features and limitations of think-aloud protocols (categorized as introspective verbal report), especially if used to investigate the concrete reading format of gapped-text.

### **1. Wstęp**

Niniejszy artykuł dotyczy metod badawczych umożliwiających zidentyfikowanie strategii wykorzystywanych przez respondentów w testach czytania. Termin *raport werbalny* odnosi się tutaj (za Matsumoto, 1993) do szeregu konkretnych metod, których cechą wspólną jest dostarczanie danych w formie przekazu językowego. W prezentowanym tekście opisane zostały jednak tylko wybrane aspekty tej bardzo obszernej kategorii, a mianowicie:

a) poszczególne rodzaje raportu werbalnego, b) problematyczność rozróżnienia pomiędzy introspekcyjnym a retrospekcyjnym raportem werbalnym w badaniach nad testowaniem czytania, c) problemy związane z trafnością raportu oraz d) cechy charakterystyczne introspekcji, a konkretnie *protokołu głośnego myślenia*, i jej zróżnicowane realizacje. Prezentowany tutaj opis ponadto w sposób szczególny odnosi się do zadań typu tekst z lukami (ang. *gapped-text*) i, w odniesieniu to tego właśnie

formatu, wskazuje na trudności związane z zastosowaniem introspekcji. Artykuł ma w dużej mierze charakter teoretycznej dyskusji wskazującej na te aspekty raportu werbalnego, które często pozostają niezauważone.

## 2. Raport werbalny intro- i retrospekcyjny

Raport werbalny polega na eksternalizacji myśli, działań lub odczuć (a w konkretnym, omawianym tu przypadku – strategii) poprzez ujęcie ich w przekaz językowy. Raport werbalny nie jest kategorią jednolitą – w jego obrębie wyróżnić można szereg zmiennych, których odmienna realizacja wpływa na zróżnicowanie konkretnych metod badawczych. I tak, w sposób najbardziej podstawowy, raport werbalny może być podzielony w oparciu o dwa czynniki: *czas realizacji* oraz *zależność pomiędzy informacją przetwarzaną i werbalizowaną* na dwie podkategorie: *raport introspekcyjny* i *retrospekcyjny* (Matsumoto, 1993). W raporcie introspekcyjnym werbalizacja przebiega równocześnie z działaniem, np. wykonywaniem zadania. W tym przypadku relacja pomiędzy informacjami przetwarzanymi i werbalizowanymi jest bezpośrednia: informacje przetwarzane są na bieżąco werbalizowane, a produktem raportu jest zapis niezmodyfikowanych informacji przetwarzanych w kolejnych momentach wykonywania zadania. Taki zapis może robić wrażenie dość chaotycznego. W badaniach introspekcyjnych respondent nie powinien jednak angażować się w nadawanie spójności swojej wypowiedzi poprzez opisywanie i interpretację podejmowanych działań. W tym przypadku interpretacja zaobserwowanych zachowań jest zadaniem badacza (Pressley i Afflerbach, 1995). Ponadto introspekcja, w odróżnieniu od retrospekcji, ma zawsze odniesienie do konkretnego wykonywanego w danym momencie działania.

W raportach retrospekcyjnych werbalizacja następuje dopiero po wykonaniu zadania. Retrospekcja polega więc na wydobyciu informacji z pamięci długotrwałej (Matsumoto, 1993), a jej produkt ma bardziej opisowy charakter. Strategie podejmowane przez respondenta są tutaj relacjonowane, nie ukazywane w sposób surowy i bezpośredni.

## 3. Intro- i retrospekcja w badaniu procesu czytania ze zrozumieniem

Raporty realizowane na bieżąco (w czasie wykonywania zadania, ang. *concurrent verbal reports*) w przypadku badań nad czytaniem mają formę tzw. *protokołu głośnego myślenia* i najczęściej klasyfikuje się je jako introspekcyjne. Czasami implikują one jednak także użycie retrospekcji<sup>6</sup> (Hosenfeld, 1977). Trudność związana z jednoznacznym zakwalifikowaniem raportu realizowanego na bieżąco jako intro- lub retrospekcyjnego wynika ze specyfiki czytania. Aby lepiej wyjaśnić tę kwestię, należy zauważyć, że raport werbalny introspekcyjny w postaci protokołu głośnego myślenia w swojej pierwotnej wersji stosowany był w badaniach nad strategiami rozwiązywania problemów (patrz

<sup>6</sup> Hosenfeld klasyfikuje protokół głośnego myślenia w formie niestymulowanej jako introspekcyjny, a w formie stymulowanej (implikującej relacjonowanie strategii po przeczytaniu fragmentu tekstu, np. jednego zdania) jako retrospekcyjny (1977).

np. Dunker, 1926), gdzie problemy były zazwyczaj stosunkowo ściśle sformułowane (polegały przeważnie na rozwiązaniu krótkiej zagadki logicznej) i nie wymagały „wczytywania dużej ilości *inputu*” (tzn. informacji prezentowanej respondentowi). „Wczytywanie” informacji przebiegało więc niejako niezależnie, a proces rozwiązywania problemu rozpoczynał się dopiero po zakończeniu czytania. Rozróżnienie w tym przypadku pomiędzy intro- i retrospekcją nie budziło więc wątpliwości. Jednak badania nad czytaniem mają nieco inny charakter. Tekst czytany wykorzystywany w badaniach z użyciem protokołu głośnego myślenia najczęściej stanowi materiał znacznie obszerniejszy niż kilkuelementowe dane liczbowe czy krótki opis problemu, co nie pozwala na wyraźne oddzielenie fazy „wczytywania” danych od pozostałych działań respondenta. Tutaj badaniu podlega ponadto proces budowania rozumienia tekstu, a proces ten przebiega równocześnie z czytaniem, nie po jego zakończeniu. Skutkuje to pojawieniem się trudności związanej z koniecznością jednoczesnego wczytywania informacji i werbalizacji strategii. W formie przyjętej przez większość badaczy (zgodnej z wytycznymi Ericssona i Simona, 1993) protokół głośnego myślenia polega na cichym czytaniu tekstu i naprzemiennie z czytaniem, w miarę na bieżąco, uzewnętrznianiu wykorzystywanych strategii. Ericsson i Simon (1993), a za nimi także inni badacze, stwierdzają, że werbalizacja po przeczytaniu krótkich fragmentów tekstu umożliwia oddawanie na bieżąco zawartości pamięci krótkotrwałej, a tym samym ma charakter introspekcyjny.

W zależności od *stopnia ingerencji badacza*, protokół głośnego myślenia może dzielić się na dwa rodzaje: *stymulowany* i *niestymulowany*. W formie mniej ingerującej – niestymulowanej – respondent czyta tekst i werbalizuje informację przetwarzaną w momencie, który sam uzna za stosowny. Całkowity brak ingerencji badacza może jednak sprzyjać pojawieniu się sytuacji, w której respondent przeczyta duży fragment tekstu lub być może nawet cały tekst nie relacjonując ani jednej strategii (Afflerbach, 1990). Prawdopodobnie z tego właśnie powodu protokół niestymulowany nie jest często wykorzystywany w badaniach nad czytaniem. Znacznie częściej stosowaną przez badaczy formą jest protokół stymulowany, w którym przypomina się respondentowi o konieczności zrelacjonowania przetwarzanych informacji za pomocą umieszczonych w tekście znaków graficznych lub/i pytań, na które respondent musi odpowiedzieć. Stosowane znaki graficzne to kropki lub gwiazdki, często w kolorze czerwonym<sup>7</sup> (patrz np. Afflerbach, 1990). Poniższy cytat prezentuje przykładowy materiał wykorzystywany w badaniu z użyciem protokołu głośnego myślenia stymulowanego znakami graficznymi i pytaniami.

Octopus is a marine animal with a soft body and eight arms, also called tentacles.\*(What do you think about this?) The word octopus \*(What do you think?) comes from two Greek words that mean eight feet.\*(How did you decide what is important to remember in this sentence? Do you have any questions that you would like to find answered in the article?) (Jimenez, García i Pearson, 1996: 97)

<sup>7</sup> Wydaje się, że czerwony kolor znaków raportu wynika z faktu powielania przez badaczy aspektów narzędzia po raz pierwszy użytego w badaniach nad czytaniem przez Jill Olshavsky (1976–77). Użycie tak jaskrawego koloru może jednak budzić pewne zastrzeżenia.

Zaprezentowany materiał jest w moim przekonaniu dość kontrowersyjny i może stanowić przykładowy punkt odniesienia w dyskusji nad raportem stymulowanym i danymi, jakich raport ten może dostarczać. W badaniu Jimeneza, Garcii i Pearsona pytania stymulujące, jak widzimy, są dłuższe niż sam tekst podstawowy. Można wręcz powiedzieć, że tekst podstawowy jest mało widoczny. Ponadto sformułowanie pytania jako „co o tym myślisz” może stymulować relację nie tyle strategii, ile różnego rodzaju innych rozważań na temat, o którym mowa w danym fragmencie (patrz pierwsze zdanie cytatu). Zastanawiające jest także zastosowanie przez autorów pytania: *what do you think?* zamiast: *what are you thinking about?*. Dwa ostatnie z cytowanych pytań stymulujących mogą także budzić zastrzeżenia: „w jaki sposób zdecydowałeś o tym, co jest ważne do zapamiętania w tym zdaniu?” i „czy masz jakieś pytania, na które chciałbyś znaleźć odpowiedź w tym artykule?” W moim przekonaniu te dwa pytania mają niewielki związek z celem badania (jakim było wydobycie strategii stosowanych przez respondentów w trakcie czytania tekstu). Pierwsze z tych pytań jest ponadto wysoko sugerujące, gdyż implikuje, że respondent uznał jakiś fragment cytowanego zdania za ważny do zapamiętania.

Analiza przedstawionego tu fragmentu, który rzeczywiście może nie być najlepszym przykładem materiału stymulującego raport, generalnie skłania do zastanowienia się nad naturą raportu stymulowanego, w szczególności w odniesieniu do użytych przez badaczy pytań stymulujących. Pytania bardzo ogólne, typu „o czym teraz myślisz” nie różnią się bowiem w działaniu od znaków graficznych, które przed rozpoczęciem wykonywania zadania zostają zdefiniowane jako miejsca, w których respondent powinien się zatrzymać i przekazać na głos swoje myśli. Natomiast pytania bardziej konkretne mogą być zawsze w jakimś kierunku sugerujące, a tym samym mogą wpływać na jakość i ilość uzyskiwanych w raporcie informacji. Jeszcze jedną problematyczną kwestią, powiązaną z trudnym w przypadku czytania rozróżnieniem pomiędzy intro- a retrospekcją, jest częstość umieszczenia znaków lub pytań stymulujących i wybór konkretnego miejsca ich umieszczenia. Większość badaczy umieszcza znaki raportu po każdym zdaniu. Czasami jednak badacze decydują się także na umieszczenie znaku raportu dodatkowo wewnątrz zdania, obok wyrazów lub sformułowań, które, w ich przekonaniu, mogą stanowić pewną trudność dla respondentów i tym samym zwiększać zapotrzebowanie na wykorzystanie pamięci roboczej. We fragmencie cytowanym powyżej, jak widzimy, autorzy uznali, że tego rodzaju trudnym dla respondentów słowem jest *octopus*, dlatego bezpośrednio po nim umieszczony został znak raportu. Jeżeli na potrzeby interpretacji sposobu działania znaków raportu przyjmiemy wytyczne Ericssona i Simona (1993) oparte na modelu przetwarzania informacji i kategoriach pamięci krótkotrwałej (plus/minus siedmioelementowej) oraz pamięci długotrwałej<sup>8</sup>, możemy zauważyć, że dość problematyczną, wysoce indywidualną i subiektywną kwestią jest decyzja

<sup>8</sup> Ericsson i Simon (1993) opisali raport werbalny w kategoriach teorii przetwarzania informacji, rozróżniając pomiędzy raportem introspekcyjnym, wymagającym od respondenta zrelacjonowania informacji znajdujących się w pamięci krótkotrwałej w poszczególnych

o miejscach umieszczenia znaków raportu. Decyzja ta, jak i jej konsekwencje polegające na nadaniu bardziej introspekcyjnego (opierającego się całkowicie na uzewnętrznianiu zawartości pamięci krótkotrwałej) lub bardziej retrospekcyjnego (wykorzystującego mimo wszystko zasoby pamięci długotrwałej) charakteru raportu jest wysoce zależna od wiedzy badacza na temat badanej grupy respondentów, ich możliwości przetwarzania informacji oraz tego, co może stanowić dla nich problem. Opierając się na modelu przetwarzania informacji, można stwierdzić, że rozróżnienie pomiędzy intro- i retrospekcją jest w wielu przypadkach sprawą dyskusyjną, a poprawne zakwalifikowanie badania jako wykorzystującego introspekcję lub retrospekcję jest w dużej mierze zależne od badacza i jego poprawnej oceny bardzo lokalnych, indywidualnych uwarunkowań realizowanego przez niego badania. Określenie potencjalnie trudnych dla respondentów miejsc w tekście, wykonywane w sposób *a priori* i w oparciu o doświadczenie badacza, rzuca jednocześnie cień na eksploracyjny charakter wielu badań z użyciem raportu werbalnego i pokazuje, że mają one także częściowo cechy podejścia dedukcyjnego.

Ze względu na kolejną zmienną – *czas realizacji* - raport werbalny *retrospekcyjny* można podzielić na realizowany bezpośrednio po wykonaniu zadania (ang. *immediate*) oraz realizowany niebezpośrednio, tzn. z pewnym opóźnieniem (ang. *delayed*) - po wykonaniu kilku zadań lub w ogóle dużo później. Jako retrospekcyjne klasyfikowane są także przez Matsumoto (1993) wszelkie badania nie dotyczące konkretnych wykonywanych wcześniej zadań, ale dotyczące całokształtu doświadczenia uczącego się. Retrospekcja może być ponadto realizowana z udziałem różnego rodzaju *wskazówek*, mających na celu przypomnienie respondentowi jego działania (może on jeszcze raz widzieć swoje odpowiedzi, oglądać zapis video swojego działania itp.).

#### 4. Trafność badań

Trafność raportu werbalnego odnosi się do możliwości uzyskania przy jego wykorzystaniu danych obrazujących rzeczywiste myśli i działania respondentów. Brak trafności raportu może mieć formę *reaktywności* (ang. *reactivity*) lub/i *braku wiarygodności* (ang. *nonveridicality*). W odniesieniu do sposobów rozwiązywania zadań, możemy powiedzieć, że raport werbalny jest reaktywny, jeżeli werbalizacja zniekształca sposób lub wydłuża czas wykonania zadania (Russo, Johnson i Stephens, 1989). O ile jednak wydłużenie czasu nie wpływa w sposób jakościowy na wnioski dotyczące stosowanych strategii, o tyle zmiana procesu powoduje, że wnioski dotyczące strategii są nietrafne. Raport jest niewiarygodny, jeżeli nie odzwierciedla poprawnie przebiegającego procesu głównego. Brak wiarygodności wiąże się z błędami wynikającymi z *ominięcia* (niezrelacjonowania pewnych strategii) lub *dodania* (zrelacjonowania procesów, które nie wystąpiły).

---

momentach wykonywania zadania oraz raportem retrospekcyjnym, polegającym na wydobyciu informacji z pamięci długotrwałej.

## 5. Trafność retrospekcji

Najpoważniejszym problemem retrospekcji jest odległość czasowa pomiędzy wykonywaniem zadania a relacjonowaniem strategii i związane z nią błędy pamięci: zapominanie, wymieszanie wiedzy wcześniejszej z wiedzą teraźniejszą (Olshavsky, 1976–77), a także fabrykacja strategii w rzeczywistości nie wykorzystywanych (Russo, Johnson i Stephens, 1989; Matsumoto, 1993). Uważa się, że metody retrospekcyjne są mniej trafne ze względu na fakt, że mogą zawierać więcej nieuzasadnionych dywagacji na temat procesów, które mogły zajść (Afflerbach, 1990). Respondenci mogą relacjonować strategie, które, jak uważają, powinni byli wykorzystywać, nie zaś te, które w rzeczywistości wykorzystywali (Matsumoto, 1993). Generalnie raport bardziej oddalony w czasie jest raportem mniej dokładnym. Garner (za Matsumoto, 1993) stwierdza eksperymentalnie, że raporty retrospekcyjne zrealizowane bezpośrednio po wykonaniu zadania i dwa dni później różnią się, i że raporty późniejsze zawierają znacznie mniej informacji. Matsumoto (1993) dodaje, że w przypadku retrospekcji, która nie odnosi się do żadnego konkretnego zadania, istnieje większe ryzyko braku trafności.

## 6. Trafność introspekcji

Za Ericssonem i Simonem, większość badaczy utrzymuje, że poprawnie przeprowadzony protokół głośnego myślenia wydłuża czas wykonania zadania, jednak nie zmienia w sposób istotny samego procesu głównego, tzn. sposobu wykonania zadania (np. Conrad, Blair i Tracy, 1999). Stwierdzenie to bywa jednak podważane przez wyniki badań innych autorów. Russo, Johnson i Stephens (1989) wykazali eksperymentalnie, że protokół głośnego myślenia jest reaktywny i wskazali na kilka potencjalnych źródeł jego reaktywności, m.in.: a) dodatkowe zapotrzebowanie na zasoby pamięci umożliwiające przetwarzanie, b) *feedback* słuchowy, c) zmianę w motywacji w kierunku dążenia do większej poprawności i asekuracyjności<sup>9</sup>. Badacze ponadto wskazują na fakt, iż respondenci mogą ulegać wpływom zaprezentowanych im przykładów w przekonaniu, że oczekuje się od nich werbalizacji podobnych do pokazanych (Pressley i Hilden, 2004; Matsumoto, 1993), dlatego też jako potencjalnie wpływającą na trafność raportu wymieniają przedstawioną respondentom instrukcję działania oraz wszelkie techniki prezentowania i modelowania sposobów tworzenia raportu. Problematiczne wydaje się także zatrzymywanie czytania poprzez znaki raportu w raporcie stymulowanym. Crain-Thoreson i in. (1997) porównali wpływ protokołu stymulowanego i niestymulowanego na wyniki czytania, stwierdzając, że prowokowanie zatrzymania się i zrelacjonowania strategii w konkretnych narzuconych przez eksperymentatorów miejscach ma negatywny wpływ na respondentów o ogólnie niższych

<sup>9</sup> Russo, Johnson i Stephens (1989) stwierdzają, że z racji tego, iż respondenci zdają sobie sprawę z publicznego występu, będą oni skłonni stosować strategie inne niż normalnie, np. wymagające większego wysiłku, ale dające większe szanse na poprawną odpowiedź.

umiejętnościach czytania, natomiast nie wpływa na wyniki respondentów o wyższych umiejętnościach<sup>10</sup> (Pressley i Hilden, 2004). Autorzy dochodzą do wniosku, iż decyzja o wyborze konkretnego typu protokołu powinna być uzależniona od relatywnego poziomu trudności tekstu: w przypadku tekstów łatwych zalecają użycie protokołu niestymulowanego, natomiast w przypadku tekstów trudnych – protokołu stymulowanego (Pressley i Hilden, 2004)<sup>11</sup>. Garner, Wagoner i Smith (1983) są bardziej krytyczni i stwierdzają, że protokół stymulowany może zniekształcać lub przerywać proces przetwarzania informacji. Według Matsumoto (1993) zniekształcenie procesów myślowych lub ich przerwanie jest bardziej prawdopodobne wtedy, gdy znaki raportu występują często lub kiedy zadanie jest bardzo trudne. Dodatkowym źródłem braku wiarygodności raportu werbalnego mogą być ograniczone umiejętności językowe respondenta, negatywnie wpływające na jakość i ilość uzyskanych danych (Matsumoto, 1993).

## **7. Wykorzystanie protokołu głośnego myślenia w badaniach nad formatem typu ‘tekst z lukami’ – dalsze rozważania na temat trafności metody**

Raport introspekcyjny, szczególnie w wersji stymulowanej, może, w moim przekonaniu, w sposób całkowicie sztuczny wywoływać użycie pewnych strategii. Stymulacja (w postaci znaków graficznych lub pytań) wymusza na przykład liniowy sposób czytania, zdanie po zdaniu i paragraf po paragrafie, z jednakowym natężeniem uwagi i w sposób jakościowo jednolity (bez zróżnicowania technik czytania np. na szybkie przeglądnięcie fragmentu, opuszczenie fragmentu, *scanning*, *skimming* itp.). Takie jednolite czytanie wynika z faktu, że, w percepcji respondenta i zgodnie z instrukcją, jaką otrzymuje, po każdym zdaniu (w takiej przeważnie odległości występują znaki raportu) musi on zrelacjonować strategie stosowane do budowania rozumienia tekstu, w tym także konkretnego zdania. Żadne ze zdań nie może więc być przez respondenta opuszczone lub pobieżne „przeskanowane”. Ponadto, raport introspekcyjny sprzyja pojawieniu się pewnych strategii, które wynikają raczej z potrzeby relacjonowania myśli, niż samego czytania tekstu, i które mogłyby się nie pojawić lub pojawić w znacznie mniejszej ilości, gdyby nie zachodziła potrzeba wykonywania czynności produktywnej – mówienia w celu relacjonowania przetwarzanych informacji. Mam tu przede wszystkim na myśli parafrazowanie, podsumowanie tekstu, sformułowanie głównej myśli, tłumaczenie na język rodzimy. Przykładem dla powyższego stwierdzenia, jakkolwiek najprawdopodobniej tylko jednym z wielu, mogłoby być badanie przeprowadzone przez Hosenfeld (1977).

<sup>10</sup>Badanie odnosi się do języka natywnego.

<sup>11</sup>Dość zaskakujący wydaje się postulat autorów w obliczu badań, jakie sami cytują. Ich badanie wydaje się bowiem sugerować, choć to dość trudne do wytłumaczenia z teoretycznego punktu widzenia, że stymulacja powinna być stosowana tam, gdzie jest ona relatywnie nieingerująca w sam proces główny, tzn. do badania tekstów relatywnie łatwych, i odwrotnie – eliminowana w sytuacji, w której wykazuje ona negatywny wpływ na czytanie, czyli w przypadku użycia tekstów trudnych.

W badaniu tym wykorzystany został raport introspekcyjny stymulowany (z zatrzymaniem po każdym zdaniu), który dostarczył danych w postaci tłumaczenia całego czytanego tekstu lub jego dużych fragmentów na język rodzimy respondentów. W moim przekonaniu tak duża ilość tłumaczenia wynika właśnie z konieczności relacjonowania przetwarzanych informacji. Jakkolwiek tłumaczenie jest czasami wykorzystywane jako strategia czytania, to jednak czytanie, nawet na dość niskim poziomie zaawansowania językowego, nie polega tylko i wyłącznie na wykonywaniu na bieżąco tłumaczenia.

Ze względu na wymieniane tu niedoskonałości standardowej formy protokołu głośnego myślenia, w przeprowadzonym przeze mnie badaniu wykorzystującym tę metodę wprowadzone zostały pewne modyfikacje (Konieczna, 2009). Badanie, o którym mowa, dotyczyło strategii wykorzystywanych przez studentów w trakcie rozwiązywania zadania testowego typu tekst z lukami (ang. *gapped-text*). Jest to format typu zamkniętego, w którym poszczególne zdania lub akapity tekstu zostały usunięte, oznaczone jako luki w tekście i zaprezentowane w formie opcji do wyboru z jedną opcją dodatkową, niepasującą do żadnej z luk. Zastosowana w badaniu modyfikacja protokołu głośnego myślenia miała na celu zredukowanie potencjalnych źródeł nietrafności, tzn.: a) uniknięcie niebezpieczeństwa czytania i nierelacjonowania strategii, charakterystycznego dla protokołu niestymulowanego oraz b) wyeliminowanie wymuszeń liniowego czytania tekstu (o których mowa w akapicie powyżej), typowych dla protokołu stymulowanego. Ponadto, protokół głośnego myślenia stymulowany po każdym zdaniu nie mógł być zastosowany ze względu na fakt, że zadanie typu tekst z lukami wymaga działania na poziomie całych paragrafów. W tym więc przypadku dość sztuczne, a także frustrujące dla respondenta mogłoby być przerywanie mu po każdym zdaniu i pytanie go o strategię, nie pozwalając na rozpoczęcie realizacji zadania. Ostatecznie więc zastosowana przeze mnie modyfikacja protokołu głośnego myślenia polegała na potraktowaniu całego zadania jako problemu do rozwiązania. Podejście takie jest szczególnie uzasadnione faktem, iż tekst z lukami, jako zadanie zamknięte, sprzyja użyciu różnego rodzaju strategii testowych, bardziej upodabniających proces wykonywania zadania do rozwiązywania problemu niż do samego czytania. Respondenci zostali więc poproszeni o wykonanie całości zadania na głos, czyli artykułowanie zarówno strategii, jak i fragmentów czytanych. Taka forma raportu pozwoliła prześledzić typowe sekwencje powrotów do poszczególnych fragmentów tekstu, ich ilość, długość fragmentów ponownie rozważanych, a także bardzo często nieliniowy sposób czytania tekstu zawartego w zadaniu i nieliniowy sposób rozwiązywania zadania. Ponadto, zastosowana modyfikacja nie wywoływała w sposób sztuczny tłumaczeń, parafrazy i podsumowań, gdyż pozwalała na werbalizację samego tekstu. Pozwoliła więc zaobserwować kiedy i w jakiej ilości tłumaczenia, parafrazy i podsumowania są w sposób rzeczywisty wykorzystywane dla celów czytania i rozwiązywania zadania, a nie wymuszonej werbalizacji. Zastosowana modyfikacja protokołu głośnego myślenia wiązała się jednak także z pewnymi utrudnieniami. Mianowicie respondenci dość często wskazywali na fakt, że „wczytywanie” informacji na głos przeszkadza w przetwarzaniu. W trakcie prowadzenia badań zauważyliśmy jednak, że ciszenie

głosu przez respondenta wykonującego zadanie powodowało, że niedogodność związana z artykulacją była znacznie mniej odczuwalna. Można więc podejrzewać, że to konieczność poprawnej artykulacji i *feedback* głosowy, zajmując część pamięci roboczej, utrudniały przetwarzanie informacji. Kolejni respondenci byli więc informowani o tym, że mogą mówić cicho, tak aby możliwie maksymalnie zredukować negatywny wpływ artykulacji na możliwości przetwarzania informacji.

Należy stwierdzić, że badania, które wykorzystują protokół głośnego myślenia w celu zidentyfikowania strategii rozwiązywania testów z zakresu czytania dostarczają dość specyficznych danych. Prezentują one bowiem rozwiązywanie zadań z zakresu czytania jako opierające się głównie na wykorzystaniu strategii typowych dla uwarunkowań testowania (tzn. strategii typu *test management* i strategii sprytu testowego, patrz Cohen i Upton, 2007; Konieczna, 2009). Relatywnie mało widoczne są strategie związane bezpośrednio z samym procesem czytania. W przeprowadzonym przeze mnie badaniu (2009), które miało charakter eksploracyjny, wyłonione strategie dotyczyły w większości kolejności czytania poszczególnych fragmentów tekstu, liczby powrotów do poszczególnych fragmentów, sposobów wyboru odpowiedzi, wykorzystania specyficznych aspektów formatu w celu ułatwienia wykonania zadania (łącznie opisanych 28 strategii). Relatywnie niewielka ilość wyłonionych strategii dotyczyła w sposób bezpośredni budowania rozumienia tekstu czytanego (użycie inferencji, podsumowanie i określenie myśli głównej danego fragmentu, tłumaczenie fragmentu, zwrócenie uwagi na elementy budujące spójność tekstu). Podobne badanie przeprowadzone wcześniej przez Cohena i Uptona (2007) dostarczyło równie ubogich danych dotyczących strategii czytania. Badanie Cohena i Uptona dotyczyło strategii stosowanych przez respondentów w formacie *drag-and-drop* testującym sprawność czytania na teście TOEFL. Autorzy zastosowali podejście dedukcyjne, sporządzając przed wykonaniem badania długą listę strategii, które, jak wnioskują, spodziewali się zidentyfikować. Ostatecznie jednak zaprezentowali tylko wyniki dotyczące najczęściej używanych przez respondentów strategii. Tak więc, spośród dwudziestu ośmiu wstępnie zdefiniowanych strategii czytania, w wynikach wspomniane są tylko trzy: a) uważne przeczytanie fragmentu tekstu, b) szybkie przeczytanie fragmentu tekstu w poszukiwaniu konkretnych informacji i c) powtórzenie, parafraza lub tłumaczenie fraz lub zdań – lub podsumowanie fragmentów – aby wspomóc rozumienie. Większość z relacjonowanych przez Cohena i Uptona strategii należy do kategorii *test-management* (17 strategii) i skupia się wokół zaobserwowanych następujących zachowań: powrotów do konkretnych pytań, kolejności czytania pytań i opcji, sposobów wyboru opcji i sprawdzania odpowiedzi. Badanie przeprowadzone przeze mnie (2009), jak i badanie Cohena i Uptona (2007) pokazują, że protokół głośnego myślenia (w formie zarówno klasycznej, jak i zmodyfikowanej) zastosowany do badania strategii wykorzystywanych w testach czytania dostarcza głównie informacji na temat sposobów radzenia sobie z testem. Taki wynik badań może świadczyć o tym, iż uwarunkowania związane z testowaniem powodują jakościową zmianę samego procesu czytania. Możliwe jest jednak także, iż sytuacja testowania po prostu sprzyja uwewnętrznianiu właśnie strategii związanych z wykonaniem testu i powoduje względną „niewidoczność”

innego rodzaju strategii. Do podobnego wniosku dochodzą Cohen i Upton (2007), którzy stwierdzają, że badany przez nich format czytania traktowany jest przez respondentów głównie w kategoriach zadania testowego, na które trzeba udzielić poprawnej odpowiedzi, nie zaś tekstu który stanowi w jakiś sposób wartościową lekturę. Biorąc jednak pod uwagę stwierdzenie, iż protokół głośnego myślenia jako metoda badawcza może sprzyjać ujawnianiu pewnego typu informacji a nieujawnianiu innych (Pressley i Afflerbach, 1995), należy także przyjąć, że możliwe jest, iż metoda ta zastosowana do badania zadań testowych z zakresu czytania daje większe szanse ujawnienia strategii typowo testowych niż innego rodzaju strategii.

## **8. Protokół głośnego myślenia w badaniach nad testowaniem czytania a kwestia czasu**

W badaniach z użyciem raportu werbalnego nie ustanawia się limitu czasowego, aby nie ograniczać werbalizacji. Istnieje jednak pewna sprzeczność pomiędzy badaniem trafności testów z użyciem raportu werbalnego (nieograniczonego czasowo) oraz faktem, iż udzielanie odpowiedzi na pytania testowe jest z natury rzeczy bardzo restrykcyjnie ograniczone czasowo. Ronald Carver zauważa że czynnik czasu ma poważny wpływ na czytanie i może w sposób istotny zniekształcać badania, w których jest on pominięty (1992). Carver stwierdza, iż osoba o wyższej motywacji będzie skłonna spędzić nad danym tekstem większą ilość czasu i włożyć więcej wysiłku w jego zrozumienie. Dlatego, w przekonaniu autora, wszelkie badania nie kontrolujące limitu czasu dostarczają zniekształconych wyników: korzystniejszych dla osób o wyższej motywacji i zaniżonych w przypadku osób słabiej zmotywowanych (Carver, 1992). Negatywne efekty związane z brakiem kontroli czasu były widoczne także w przeprowadzonym przeze mnie badaniu (patrz Konieczna, 2009) na przykładzie ucznia, który wykorzystał dwa razy więcej czasu niż średnio inni respondenci i odpowiedział bardzo dobrze na zadanie, pomimo faktu, iż jakościowo strategię przez niego wykorzystywane były typowe dla słabszych respondentów.

## **9. Wnioski**

W niniejszym artykule opisane zostały cechy charakterystyczne poszczególnych rodzajów raportu werbalnego, ze szczególnym uwzględnieniem protokołu głośnego myślenia i sposobów jego wykorzystania na potrzeby badania strategii stosowanych w testach czytania. Artykuł ukazał dość liczne problemy związane z wykorzystaniem introspekcji w badaniach nad testowaniem czytania. Zaprezentował także różnicowany charakter danych możliwych do uzyskania przy wykorzystaniu różnych typów raportu werbalnego.

Hosenfeld klasyfikuje protokół głośnego myślenia w formie niestymulowanej jako introspekcyjny, a w formie stymulowanej (implikującej relacjonowanie strategii po przeczytaniu fragmentu tekstu, np. jednego zdania) jako retrospekcyjny (1977).

Wydaje się, że czerwony kolor znaków raportu wynika z faktu powielania przez badaczy aspektów narzędzia po raz pierwszy użytego w badaniach nad czytaniem przez Jill Olshavsky (1976–77). Użycie tak jaskrawego koloru może jednak budzić pewne zastrzeżenia.

Ericsson i Simon (1993) opisali raport werbalny w kategoriach teorii przetwarzania informacji, rozróżniając pomiędzy raportem introspekcyjnym, wymagającym od respondenta zrelacjonowania informacji znajdujących się w pamięci krótkotrwałej w poszczególnych momentach wykonywania zadania oraz raportem retrospekcyjnym, polegającym na wydobyciu informacji z pamięci długotrwałej.

Russo, Johnson i Stephens (1989) stwierdzają, że z racji tego, iż respondenci zdają sobie sprawę z publicznego występu, będą oni skłonni stosować strategie inne niż normalnie, np. wymagające większego wysiłku, ale dające większe szanse na poprawną odpowiedź.

Badanie odnosi się do języka natywnego.

Dość zaskakujący wydaje się postulat autorów w obliczu badań, jakie sami cytują. Ich badanie wydaje się bowiem sugerować, choć to dość trudne do wytłumaczenia z teoretycznego punktu widzenia, że stymulacja powinna być stosowana tam, gdzie jest ona relatywnie nieingerująca w sam proces główny, tzn. do badania tekstów relatywnie łatwych, i odwrotnie – eliminowana w sytuacji, w której wykazuje ona negatywny wpływ na czytanie, czyli w przypadku użycia tekstów trudnych.

## BIBLIOGRAFIA

- Afflerbach, P.P. 1990. « The influence of prior knowledge on expert readers' main idea construction strategies ». *Reading Research Quarterly* 25 : 31–46.
- Carver, R. 1992. « Effect of prediction activities, priori knowledge, and text type upon Mount comprehended: using rauding theory to critique schema theory research ». *Reading Research Quarterly* 27 : 164–174.
- Cohen, A.D., Upton, T.A. 2007. « 'I want to go back to the text': response strategies on the reading subtest of the new TOEFL ». *Language Testing* 24 : 209–250.
- Conrad, F., Blair, J., Tracy, E. 1999. « Verbal reports are data! A theoretical approach to cognitive interviews ». Proceedings of the Federal Committee on Statistical Methodology Research Conference, [www.bls.gov/osmr/abstract/st/st990240.htm](http://www.bls.gov/osmr/abstract/st/st990240.htm) DW 19.07.2009.
- Ericsson, K. A., Simon, H. A. 1993. *Protocol analysis: Verbal reports as data*. Cambridge : Bradford books/MIT Press.
- Garner, R., Wagoner, S., Smith, T. 1983. « Externalizing question-answering strategies of good and poor comprehenders ». *Reading Research Quarterly* 18 : 439–447.
- Hosenfeld, C. 1977. « A preliminary investigation of the reading strategies of successful and unsuccessful second language learners ». *System* 5 : 110–123.

- Jimenez, R. T., Garcia, G. E., Pearson, P. D. 1996. « The reading strategies of bilingual Latina/o students who are successful English readers: opportunities and obstacles ». *Reading Research Quarterly* 31 : 90–112.
- Koniczna, A. 2009. *The use of strategies in gapped-text reading task*, referat konferencyjny, 3<sup>rd</sup> International May Conference on English Studies, Kraków [w druku].
- Matsumoto, K. 1993. « Verbal report data and introspective methods in second language research: state of the art ». *RELC Journal* 24 : 32–60.
- Olshavsky, J. E. 1976–1977. « Reading as problem solving: an investigation of strategies ». *Reading Research Quarterly* 12 : 654–674.
- Pressley, M., Afflerbach, P. 1995. *Verbal protocols of reading: the nature of constructively responsive reading*. Hillsdale : Lawrence Erlbaum Associates.
- Pressley, M., Hilden, K. 2004. « Verbal protocols of reading » (w) *Literacy research methodologies* (red. N. K. Duke i M. H. Mallette). New York : The Guilford Press.
- Russo, J. E., Johnson, J. J., Stephens, D. L. 1989. « The validity of verbal protocols ». *Memory and Cognition* 17 : 759–769.