

## Koneksjonistyczne modele mechanizmów rozwoju poznawczego

Adam Putko<sup>1</sup>

Instytut Psychologii, Uniwersytet im. A. Mickiewicza  
Szkoła Wyższa Psychologii Społecznej

### CONNECTIONIST MODELS OF MECHANISMS OF COGNITIVE DEVELOPMENT

**Abstract.** This paper presents connectionist models of cognitive development, which yield stage-like behavioural properties, but are based on static or generative neural networks. These networks differ in learning rules that they use. Static networks use only one kind of mechanism of change in information processing that consists in quantitative changes of units connectivity. Generative networks can use more powerful kind of mechanism of change that consists in quantitative as well qualitative changes in net structure by recruitment of new units. It is discussed how connectionist approach can provide insight into the role that an environmental input and learning constraints play in the cognitive development. Simulation models of the acquisition of English past tense and the modular model of the development of seriation are described that help illustrate the problem of quantitative and qualitative mechanisms of change. The article concludes that different types of developmental changes can result from the interaction between the specific requirements of task (e.g. conflicting types of input-output mappings) and learning constraints.

Znaczenie podejścia koneksjonistycznego do badań nad rozwojem poznawczym stało się przedmiotem wielu analiz zarówno teoretycznych (por. Bates, Elman, 1993; Elman i in. 1996; Mareschal, Shultz, 1996; McClelland, 1995; Plunkett, Sinha, 1992), jak i empirycznych (por. Mareschal, Shultz, 1999; McClelland, 1989; Quinn, Johnson, 1997). Jest ono obecnie – obok podejścia opartego na teorii systemów dynamicznych (por. Thelen, Smith, 1994; van Geert, 1994; van der Maas, Hopkins, 1998) – jednym z najbardziej prężnych nurtów w modelowaniu rozwoju poznawczego. Wyraża się ono przede wszystkim w tworzeniu modeli mechanizmów rozwoju, odwołujących się do zasad przetwarzania informacji przez sieci neuronowe, a stąd wystarczająco explicytnych, aby mogły być poddawane bezpośredniemu sprawdzeniu na drodze symulacji i porównywania uzyskiwanych wyników z danymi empirycznymi.

Celem niniejszego artykułu jest ukazanie możliwości i ograniczeń koneksjonistycznego podejścia do modelowania mechanizmów rozwoju poznawczego. Problem mechanizmów rozwoju pozostaje jednym z istotnych zagadnień psychologii rozwojowej i wiąże się z pytaniem, w jakim stopniu u podstaw behawioralnych zmian rozwojowych leżą ilościowe lub jakościowe zmiany w sposobie przetwarzania informacji. Pytanie to jest częścią szerszego problemu, dotyczącego natury zmian rozwojowych. Spośród dwóch istniejących w tej dziedzinie stanowisk jedno zakłada, że zmiany mają charakter nieciągły i stadialny (por. Piaget, 1972; Case, 1985; van Geert, 1994), a drugie – że rozwój przebiega w sposób ciągły i łagodny (por. Siegler, 1995; 1996). Równie ważnym problemem, na którego rozwiązanie rzutują wyniki koneksjonistycznych symulacji, pozostaje pytanie o źródło zmian rozwojowych: w jakim stopniu zmiany są wynikiem wewnętrznych ograniczeń związanych z procesem uczenia, a w jakim stopniu wynikiem zewnętrznych względem organizmu czynników, takich jak wymagania, jakie stawia zadanie, i specyficzne cechy informacji, które są przetwarzane.

### PODEJŚCIE KONEKSJONISTYCZNE DO MODELOWANIA ROZWOJU

Charakterystyczną cechą podejścia koneksjonistycznego, ważną z punktu widzenia modelowania rozwoju poznawczego, jest to, że przejawiane przez system umiejętności i wiedza są wynikiem procesu uczenia się. Model koneksjonistyczny jest od strony swej struktury siecią neuropodobnych jednostek, przetwarzających informacje w sposób równoległy. Istnieje wiele metod uczenia, z których najczęściej wykorzystywane jest uczenie nadzorowane. Polega ono na podawaniu na wejście sieci przykładów uczących i porównywaniu udzielanych na wyjściu odpowiedzi z odpowiedziami pożądanymi. Po wyznaczeniu wielkości błędu (rozbieżności między rzeczywistym a pożądanym zachowaniem) dokonywane są zmiany w wagach połączeń między jednostkami według odpowiedniego dla struktury sieci algorytmu. Zmiany te zwiększają prawdopodobieństwo aktywacji jednostek pożądaných i zmniejszają prawdopodobieństwo aktywacji jednostek niepożądaných. Cykl uczenia powtarzany jest tak długo, aż zachowanie systemu osiągnie zadowalający poziom.

1 Adres do korespondencji: Instytut Psychologii UAM, ul. Szamarzewskiego 89, 61-568 Poznań; e-mail: putko@amu.edu.pl

ADAM PUTKO

Wiedza, jaką zdobywa sieć neuronowa w wyniku uczenia się, zakodowana jest w wagach połączeń między jednostkami. To właśnie wagi połączeń determinują sposób, w jaki obrazy wejściowe są odwzorowywane na obrazy wyjściowe. Wiedza zakodowana w wagach połączeń między jednostkami może być traktowana jako odpowiednik reprezentacji implicytnych, natomiast wzorce aktywacji tzw. jednostek ukrytych – jako odpowiednik reprezentacji explicytnych. Zmiany wiedzy posiadanej przez system można opisać jako ścieżkę w przestrzeni możliwych połączeń synaptycznych, tzn. jako transformację wektora wag w  $n$ -wymiarowej przestrzeni wag, gdzie  $n$  oznacza liczbę jednostek ukrytych.

Pomimo znacznej „plastyczności” sieci neuronowej, której układ połączeń kształtowany jest w wyniku uczenia, nie każda struktura sieci jest adekwatna do rozwiązywania określonego zadania. Istotne znaczenie ma jej architektura – liczba warstw jednostek (sieci jednowarstwowe vs. wielowarstwowe), kierunek przepływu sygnałów (sieci jednokierunkowe vs. sieci rekurencyjne) czy też rodzaj funkcji aktywacji charakteryzującej jednostki (funkcja liniowa vs. nieliniowa). Sieci o architekturze jednowarstwowej zdolne są przykładowo do rozwiązywania jedynie problemów tzw. separowalnych liniowo. Nie potrafią one rozwiązać problemu XOR (alternatywy wykluczającej „albo-albo”), do którego rozwiązania zdolne są sieci mające warstwę ukrytą, pośredniczącą między warstwą wejściową i wyjściową. Również rodzaj funkcji aktywacji jednostek może decydować o stopniu dopasowania sieci do rozwiązywania określonego zadania. Na przykład Jacobs, Jordan i Barto (1991) wykazali, że jeżeli jedna sieć składa się z jednostek o liniowej funkcji aktywacji, a druga z jednostek o nieliniowej funkcji aktywacji, wówczas pierwsza nauczy się rozwiązywać zadanie typu „gdzie” – związane z lokalizowaniem obiektu w przestrzeni, a druga zadanie typu „co” – łączące się z identyfikacją obiektu. „Preferencja” zadania przestrzennego przez sieć o liniowej funkcji aktywacji wynika z lepszego dopasowania dynamiki zmian jej stanów reprezentacyjnych do liniowej struktury przestrzeni. Przykład ten pokazuje, że natura problemu i jego wymagania komputacyjne narzucają ograniczenia na sposób jego rozwiązania oraz na systemy, które mogą temu rozwiązaniu podolać.

Z punktu widzenia mechanizmów uczenia wykorzystywanych w koneksjonizmie ważne jest odróżnienie sieci statycznych o nie zmieniającej się w trakcie uczenia architekturze od sieci generatywnych – o zmieniającej się architekturze. W sieciach statycznych mechanizm uczenia ogranicza się do ilościowych zmian w parametrach połączeń między jednostkami, natomiast w sieciach generatywnych uczenie może wiązać się dodatkowo ze zmianą liczby jednostek i połączeń (zmiany strukturalne). Przykładem sieci wykorzystujących mechanizm strukturalnych zmian progresywnych są sieci kaskadowej korelacji (Fahlman, Lebiere, 1990), w których proces uczenia wiąże się z „rekrutowaniem” nowych jednostek. Z kolei mechanizm strukturalnych zmian regresywnych, związanych z eliminacją jednostek i połączeń, wykorzystywany jest w sieciach meiotycznych (Hanson, 1990).

W związku z rozróżnieniem na sieci statyczne i generatywne pojawia się pytanie dotyczące natury jakościowych zmian w zachowaniu, związanych z przechodzeniem do nowych stadiów rozwojowych. W jakim stopniu u podłoża tych behawioralnych zmian leżą jakościowe i ilościowe zmiany w mechanizmach przetwarzania informacji? Aby odpowiedzieć na to pytanie, porównano modele rozwoju poznawczego budowane na podstawie statycznych oraz generatywnych sieci kaskadowej korelacji. Architektura sieci statycznej, która ma rozwiązać określone zadanie, musi zostać zaprojektowana przed rozpoczęciem procesu uczenia. Zdolność sieci statycznych do wytwarzania adekwatnych reprezentacji zadania okazała się wystarczająca w modelowaniu wielu zjawisk rozwojowych (np. Elman i in. 1996; McClelland, 1995; Plunkett, Sinha, 1992; Quinn, Johnson, 1997). Jednakże zbiór reprezentacji, które mogą zostać przez taką sieć wytworzone, jest od początku ograniczony przez jej architekturę (por. Cybenko, 1989; Mareschal, Shultz, 1996; Quartz, 1993). W modelowaniu rozwoju poznawczego istotna wydaje się możliwość zmiany zasobów reprezentacyjnych i komputacyjnych, którymi dysponuje system. Generatywne sieci kaskadowej korelacji mają takie możliwości, co sprawia, że budowane na ich podstawie modele rozwoju poznawczego mogą trafniej oddawać rzeczywisty przebieg zmian rozwojowych niż modele budowane na podstawie sieci statycznych. Na przykład statyczne sieci dysponujące zbyt małymi zasobami reprezentacyjnymi często przejawiają trudności w osiągnięciu końcowego stadium rozwojowego (stadium multiplikatywne) w rozwiązywaniu problemu „wagi” (Siegler, 1981). Z kolei sieci dysponujące od początku dużymi zasobami reprezentacyjnymi zbyt szybko osiągają końcowe stadium rozwojowe, nie przejawiając cech charakterystycznych dla stadiów pośrednich (addytywnych). Podstawową zaletą modelu generatywnego jest zatem możliwość sprawdzenia, na ile obserwowane zmiany rozwojowe wymagają wzrostu zasobów reprezentacyjnych i komputacyjnych.

W dalszej części artykułu są przedstawione koneksjonistyczne modele rozwoju językowego – zbudowane na podstawie statycznych sieci neuronowych – oraz model rozwoju szeregowania, zbudowany na podstawie generatywnej sieci neuronowej. Wymienione wyżej zdolności należą do tych dziedzin, które przez długi czas uważane były za bastion klasycznej, symbolicznej teorii komputacyjnej (por. Fodor, 1981), zakładającej, że istotą procesów umysłowych jest manipulowanie explicytnie reprezentowanymi symbolami.

## MODELE ROZWOJU JĘZYKOWEGO

Model Rumelharta i McClellanda (1986; 1987) był pierwszą poważną próbą zastosowania podejścia PDP (*parallel-distributed processing*) do wyjaśniania procesów kierujących przyswajaniem języka. Ogólnym celem symulacji było ukazanie możliwości zbudowania modelu przyswajania języka przejawiającego istotne aspekty rozwojowe, a jednocześnie nie opierającego się na mechanizmach explicytnie reprezentowanych reguł i dyskretnych symboli.

Przedmiotem symulacji były fazy uczenia się przez dzieci form czasu przeszłego w języku angielskim. W rozwoju dzieci angielskojęzycznych występuje faza, w której pojawiają się błędne formy czasu przeszłego czasowników nieregularnych, takie jak *goed* lub *sitted*. Określa się je błędami regularyzacji i traktuje jako wynik nadmiernego uogólniania zasady tworzenia form czasu przeszłego dla czasowników regularnych. Tego rodzaju błędy pojawiają się zazwyczaj po fazie, w której dzieci posługiwały się w sposób poprawny formami czasu przeszłego czasowników nieregularnych. W toku dalszego rozwoju błędy regularyzacji zanikają i dzieci opanowują prawidłowe formy czasowników regularnych i nieregularnych (por. Bybee, Slobin, 1982; Kuczaj, 1977; Marchman, 1984).

Taki przebieg procesu uczenia, w którym po początkowych postępach występuje regres, a następnie poprawa, traktowany jest w koncepcjach reprezentujących podejście symboliczne jako wynik reorganizacji reguł i reprezentacji (por. Karmiloff-Smith, 1986; Pinker, Prince, 1988) – proces wyzwalany przede wszystkim przez wewnętrzne czynniki, a nie przez własności, jakimi charakteryzują się dane wejściowe. W modelu Rumelharta i McClellanda zaproponowano inne podejście do wyjaśnienia zmian rozwojowych, które zakłada, że zmiany te wywoływane są cechami środowiska językowego.

Od strony funkcjonalnej architektura modelu odpowiadała sieci jednowarstwowej. Zachowanie modelu w trakcie uczenia, opisane za pomocą tzw. krzywej uczenia się, było podobne pod pewnymi względami do zachowania dzieci uczących się form czasu przeszłego w języku angielskim. Przede wszystkim po początkowych postępach w uczeniu obniżał się poziom wykonania dla czasowników nieregularnych związany z tendencją do ich regularyzacji. Co więcej, przebieg uczenia się różnych klas czasowników nieregularnych – rodzaj błędów i kolejność, w jakiej się pojawiały – odpowiadał przebiegowi uczenia u dzieci. Na przykład błędne formy *ated* pojawiały się później niż formy typu *eated*, co jest zgodne z obserwacjami rozwojowymi (por. Kuczaj, 1977). Najważniejszy zarzut, który postawiono modelowi Rumelharta i McClellanda, dotyczył nieadekwatnego reprezentowania zbioru danych wejściowych, na podstawie którego wykrywane są przez dzieci prawidłowości systemu czasu przeszłego (por. Pinker, Prince, 1988). W początkowej fazie zbiór uczący był ograniczony do niewielkiej liczby najczęściej używanych w języku angielskim czasowników, z których większość stanowiły czasowniki nieregularne. Następnie zbiór ten poszerzono o większą liczbę czasowników, wśród których przeważały czasowniki regularne. Właśnie w tej fazie pojawiał się regres w uczeniu się czasowników nieregularnych. Pinker i Prince (1988) zauważyli, że charakteryzująca model krzywa uczenia „U” jest wynikiem nieciągłości w wielkości i strukturze danych wejściowych. Dane rozwojowe (por. Brown, 1973) nie wskazują, aby zasób czasowników regularnych w słownictwie dzieci ulegał istotnej zmianie tuż przed pojawieniem się fazy regularyzacji, a zatem założenie w modelu Rumelharta i McClellanda jest nieadekwatne.

W odpowiedzi na te zarzuty Plunkett i Marchman (1991) oraz Plunkett, Marchman i Knudsen (1991) wykazali, że podstawowe własności rozwojowe modelu koneksjonistycznego mogą zostać zachowane, nawet jeżeli zrezygnuje się z wątpliwego założenia dotyczącego struktury danych wejściowych. Stwierdzili oni, że zadanie, jakim jest wyuczenie się systemu czasu przeszłego języka angielskiego, przekracza zdolności komputacyjne jednowarstwowej sieci i że konieczne jest przyjęcie modelu, który oprócz warstwy wejściowej i wyjściowej, ma tzw. warstwę jednostek ukrytych. Sieci mające warstwę ukrytą zdolne są do wykrywania abstrakcyjnych zależności istniejących w zbiorze danych wejściowych.

Symulacje przeprowadzone na podstawie tego modelu dowiodły, że krzywa uczenia się „U” pojawia się także wówczas, gdy wielkość zbioru uczącego pozostaje stała, a więc gdy zachowany zostaje warunek ciągłości w wielkości i strukturze danych wejściowych. Czynnikiem odpowiedzialnym za spadek poziomu poprawnych reakcji w odniesieniu do pewnych czasowników jest występowanie w zbiorze uczącym konfliktowych typów odwzorowań form podstawowych na formy czasu przeszłego. W języku angielskim takie konfliktowe typy odwzorowań związane są z czasownikami regularnymi i nieregularnymi. Wśród tych ostatnich występuje wiele różnych klas odwzorowań formy podstawowej na formę czasu przeszłego (np. odwzorowania tożsamościowe *hit-hit*, arbitralne *go-went* czy też polegające na wymianie samogłoski *come-came*). Krzywa uczenia „U” jest wynikiem dążenia sieci do znalezienia ogólnego rozwiązania dla konfliktowych typów odwzorowań przez utworzenie takiej konfiguracji wag połączeń między jednostkami, która jest zgodna ze wszystkimi typami odwzorowań.

Symulacje przeprowadzone przez Plunketta i Marchmana (1991) ujawniły również, że nie wszystkie czasowniki

ADAM PUTKO

należące do tej samej klasy odwzorowań przechodzą przez fazę regularyzacji w tym samym czasie. Efektem tego jest występowanie mikrokrzywych uczenia, pojawiających się wielokrotnie w odniesieniu do pojedynczych czasowników. Stwierdzone w symulacji koneksjonistycznej cechy rozwoju językowego znajdują potwierdzenie w analizach i obserwacjach rozwojowych. Lokalny, a nie ogólny charakter nadmiernej regularyzacji potwierdza analiza przeprowadzona przez Marcusa i współautorów (1992) na podstawie bazy danych językowych Childes (MacWhinney, 1991).

Kolejnym ważnym wynikiem symulacji (Plunkett, Marchman, 1991) było odkrycie efektu masy krytycznej w procesie uczenia. W sytuacji, gdy zbiór uczący liczył 20 czasowników, uczenie nie prowadziło do trafnych uogólnień w konfrontacji z nowymi czasownikami. Dopiero zwiększenie liczebności zbioru uczącego do ok. 40-50 elementów spowodowało podwyższenie poziomu poprawnych uogólnień. Autorzy modelu wyjaśniają ten efekt tym, że niewielka liczba przypadków w zbiorze uczącym prowadziła do wytworzenia izolowanych reprezentacji par bodziec-reakcja, co przypomina proces tzw. uczenia się na pamięć. Natomiast większa liczba przypadków w zbiorze uczącym prowadzi do wytworzenia bardziej systematycznych reprezentacji. Wynik ten wskazuje, że proces wydobywania pamięciowego i uogólnień są w sieci neuronowej realizowane przez jeden mechanizm, co może stanowić również wyjaśnienie konstruktywnego charakteru pamięci. Występowanie efektu masy krytycznej w uczeniu się języka potwierdzają wyniki analizy Marchman i Bates (1994), opartej na danych językowych pochodzących od 1000 dzieci, które wskazują, że występuje silna dodatnia korelacja między zasobem czasowników w słownictwie dziecka a tendencją do nadmiernej regularyzacji.

Interesujące wnioski wypływają z analizy wpływu częstości czasowników na przebieg uczenia się ich form czasu przeszłego (Plunkett i Marchman, 1991). Wyróżniono dwa rodzaje częstości – częstość: (1) typu, czyli liczbę czasowników należących do określonej klasy formowania czasu przeszłego, oraz (2) częstość egzemplarzową, czyli względną częstość, z jaką dany czasownik występuje w mowie. Czasowniki nieregularne w języku angielskim cechują się niską częstością typu, ale wysoką częstością egzemplarzową, natomiast czasowniki regularne – wysoką częstością typu oraz znacznie zróżnicowaną częstością egzemplarzową. Symulacje wykazały, że charakterystyczne cechy odwzorowań czasowników o dużej częstości typu niejako „przebiekają” do klas czasowników o niskiej częstości typu, prowadząc do pojawiania się błędów. Czynnikiem, który może osłabiać tego rodzaju interferencje, jest wysoka częstość egzemplarzowa, która ma miejsce w przypadku niektórych czasowników nieregularnych.

Podsumowując, przedstawione wyżej koneksjonistyczne modele rozwoju językowego podważają tradycyjny pogląd, iż przyswajanie języka polega na tworzeniu i reorganizowaniu explicity reprezentowanych reguł, łącznie z odrębnymi reprezentacjami wyjątków od tych reguł. Procesy reorganizacji, które leżą u podstaw przyswajania regularnych i nieregularnych form, są w tych modelach wynikiem działania jednolitego mechanizmu, nie rozróżniającego explicity form regularnych od nieregularnych. Wnioski, jakie płyną z przedstawionych wyżej wyników symulacji, sięgają jednak, jak się zdaje, znacznie dalej. Wskazują one, że u podstaw niektórych jakościowych zmian w zachowaniu może leżeć mechanizm jedynie ilościowych, a nie jakościowych zmian w sposobie przetwarzania informacji. Modele te sugerują również, że w powstawaniu określonych zmian rozwojowych (jakościowych vs. ilościowych) istotną rolę może odgrywać interakcja między zewnętrznymi w stosunku do organizmu czynnikami, takimi jak wymagania zadania (np. konfliktowe typy odwzorowań wejścia na wyjście), a czynnikami wewnętrznymi, związanymi z ograniczeniami w mechanizmach uczenia (np. brak możliwości pojawiania się strukturalnych zmian w sposobie przetwarzania informacji).

## MODEL ROZWOJU SZEREGOWANIA

Szeregowanie należało do podstawowych zadań wykorzystywanych przez Piageta (Piaget 1965; Piaget, Inhelder, 1973) w badaniach nad rozwojem myślenia u dzieci. Znanych jest kilka modeli szeregowania opartych na podejściu symbolicznym (por. np. Baylor i in. 1973; Young, 1976), w których rozwój szeregowania jest wynikiem przyswajania zbioru procedur i uczenia się, jak i kiedy je stosować. Modele symboliczne wykorzystują zazwyczaj explicity reprezentowane reguły typu „jeżeli-to”, które określają działania i warunki ich uruchamiania. Na przykład reguła implementująca jedno ze stadiów szeregowania może brzmieć: „Jeżeli szereg elementów jest nieuporządkowany i celem jest uporządkowanie go w kolejności od najmniejszego do największego, to należy przesunąć najmniejszy nieuporządkowany element na właściwą pozycję”.

Pierwszym modelem, który reprezentuje podejście koneksjonistyczne do problemu rozwoju szeregowania, jest model Mareschala i Shultza (1999). Model ten wykorzystuje algorytm kaskadowej korelacji, umożliwiając konstruowanie topologii sieci w trakcie procesu uczenia. Jest on częścią szerszego projektu badawczego, którego celem jest utworzenie modeli rozwoju poznawczego dotyczących różnych dziedzin na podstawie jednego ogólnego mechanizmu uczenia (Shultz i in. 1995). Ten sam algorytm uczenia został z powodzeniem zastosowany m.in. w modelu rozwiązywania problemu „wagi” (Shultz, Schmidt, 1991), a także rozwoju rozumienia wzajemnej

## KONEKSJONISTYCZNE MODELE MECHANIZMÓW ROZWOJU POZNAWCZEGO

zależności między prędkością, czasem i odległością (Buckingham, Shultz, 1994).

### ALGORYTM KASKADOWEJ KORELACJI

Algorytm kaskadowej korelacji umożliwia konstruowanie topologii sieci w trakcie uczenia, a nie przed jego rozpoczęciem, jak ma to miejsce w tzw. statycznych sieciach o nie zmieniającej się architekturze. Sieci kaskadowej korelacji mogą zwiększać swój potencjał reprezentacyjny i komputacyjny przez dołączanie nowych jednostek ukrytych, co jest zgodne z konstruktywistycznym podejściem do rozwoju (Quartz, 1993; Quartz, Sejnowski, 1997).

Uczenie według algorytmu kaskadowej korelacji (Fahlman, Lebiere, 1990) polega na „rekrutowaniu” do sieci nowych jednostek ukrytych i modyfikowaniu wag połączeń (por. rys. 1). Proces uczenia rozpoczyna się od sieci o minimalnej topologii – złożonej z jednej warstwy wejściowej połączonej z jedną warstwą wyjściową – i polega na modyfikowaniu wag połączeń między jednostkami w taki sposób, aby zminimalizować błędne odpowiedzi, jakich udziela warstwa wyjściowa. Jest to sposób uczenia typowy dla większości sieci neuronowych.

Rys. 1. Fazy uczenia według algorytmu kaskadowej korelacji: (a) początkowa faza uczenia; (b) faza uczenia z „kandydatami”; (c) kolejna faza uczenia; (d) faza uczenia z podłączonymi dwiema jednostkami ukrytymi (linią przerywaną zaznaczono połączenia modyfikowalne, linią ciągłą – połączenia zamrożone)

Istotą algorytmu kaskadowej korelacji jest to, że ten typowy sposób uczenia ulega zmianie wówczas, gdy spełniony jest jeden z dwóch warunków: (1) dalsze uczenie nie prowadzi do znaczącej redukcji błędów – co odpowiada sytuacji, w której wymagania zadania są prawdopodobnie zbyt złożone w stosunku do możliwości sieci; (2) uczenie nie zakończyło się osiągnięciem zadowolającego poziomu wykonania w pewnej ustalonej liczbie cykli – co z kolei odpowiada sytuacji, w której zachowanie jest nieefektywne. Zmiana w sposobie uczenia polega na zwiększeniu zasobów reprezentacyjnych sieci przez dołączenie do niej nowej jednostki. Początkowo do sieci jest podłączanych kilka nowych jednostek ukrytych, tzw. kandydatów. Są one połączone ze wszystkimi jednostkami sieci, z wyjątkiem jednostek wyjściowych, a wagi ich połączeń są przypadkowe. Następnie połączenia jednostek „kandydujących” są modyfikowane w taki sposób, aby zmaksymalizować absolutną wielkość kowariancji aktywacji jednostki „kandydującej” i błędów jednostek „wyjściowych”. Faza „rekrutacji” kończy się wówczas, gdy kowariancja nie ulega dalszemu polepszeniu albo proces uczenia przekroczył wcześniej określoną liczbę cykli. Jednostka „kandydat” z najwyższą absolutną kowariancją zostaje podłączona do wszystkich jednostek „wyjściowych”, a wagi jej połączeń wejściowych zostają „zamrożone”. Pozostałe jednostki „kandydujące” zostają odrzucone.

Algorytm kaskadowej korelacji odzwierciedla w znacznym stopniu zmiany, jakie zachodzą na poziomie neuronalnym mózgu w wyniku uczenia się i dojrzewania: synaptogenezę i eliminację połączeń między neuronami (Changeux, Dehaene, 1989; Dawson, Fischer, 1994; Quartz, Sejnowski, 1997). Jednakże algorytm ten może być również traktowany jako odpowiednik psychologicznych mechanizmów, związanych z przebudową struktur wiedzy. Zwiększenie zasobów reprezentacyjnych przez dołączenie nowych jednostek ukrytych może odpowiadać procesowi re-reprezentowania wiedzy (Karmiloff-Smith, 1992), w wyniku którego dotąd posiadana wiedza jest reprezentowana w nowy sposób. Ustawienie jednostek ukrytych w kaskadę, w której wszystkie następne jednostki otrzymują sygnały od poprzednich jednostek, może stanowić również neuronalny korelat mechanizmu hierarchicznej integracji wiedzy w trakcie rozwoju.

### SYMULACJA

Piagetowskie zadanie badające rozwój szeregowania polega na uporządkowaniu zbioru elementów według wielkości (Inhelder, Piaget, 1969; Piaget, 1965). W rozwoju szeregowania wyróżnia się cztery stadia. W stadium 1 (ok. 4. roku życia) dziecko nie czyni starań, aby uporządkować elementy, i pozostawia je w tej samej kolejności albo przestawia w sposób przypadkowy. W stadium 2 (ok. 5. roku życia) dziecko nie potrafi wprawdzie uporządkować całego zbioru, ale umie zestawiać ze sobą elementy według lokalnych absolutnych jakości, takich jak „duży” i „mały”. Prowadzi to do powstawania szeregów nieskoordynowanych par elementów (złożonych z jednego dużego i jednego małego elementu), nieskoordynowanych trójek (element duży, średni i mały), szeregowania opartego na prawidłowym ustawieniu tylko górnych części elementów, szeregowania dachowatego (w którym górne części podnoszą się, a potem opadają lub odwrotnie) oraz szeregowania, w którym pierwsze elementy zostały prawidłowo uporządkowane, lecz kolejnych dziecko nie potrafiło już uporządkować. W stadium 3, zwanym empirycznym (ok. 6. roku życia), dziecku udaje się uszeregować elementy, lecz tylko metodą prób i błędów. Wreszcie w 4 stadium operacyjnego szeregowania (ok. 7. roku życia) dziecko szereguje systematycznie,

ADAM PUTKO

wybierając zawsze najkrótszy (lub najdłuższy) element spośród tych, które jeszcze nie zostały uporządkowane i umieszcza go we właściwym miejscu w szeregu.

Model Mareschala i Shultza (1999) zakłada, że chociaż szeregowanie jest związane z przeprowadzaniem szeregowo uporządkowanych działań, każde z nich może być wynikiem odrębnych procesów przebiegających równolegle: jednego odpowiedzialnego za wyznaczenie elementu, który ma zostać przeniesiony, i drugiego, odpowiedzialnego za miejsce jego przeniesienia. Zgodnie z tym założeniem model składa się z dwu modułów, przetwarzających te same dane wejściowe, lecz działających niezależnie od siebie (por. rys. 2). Zadaniem jednego modułu jest wyznaczenie elementu, który powinien zostać przeniesiony (moduł „który”), a zadaniem drugiego – miejsca jego przeniesienia (moduł „gdzie”). Każdy z modułów poddawany jest niezależnemu uczeniu.

Rys. 2. Schemat koneksjonistycznego modelu szeregowania Mareschala i Shultza (1999) (kolorem szarym zaznaczono jednostki ukryte)

Powodem zastosowania modularnej architektury były niezbyt udane próby utworzenia niemodularnego modelu szeregowania. Zastosowanie rozwiązania modularnego pozwoliło na lepszą funkcjonalną dekompozycję zadania, dzięki czemu elementarne składniki zadania stały się łatwiejsze do wyuczenia. Założenie dotyczące modularnego charakteru mechanizmu szeregowania wymaga weryfikacji empirycznej w badaniach psychologicznych. Podobne modularne rozwiązanie występuje w systemie wzrokowym, gdzie informacja o kształcie obiektu jest przetwarzana niezależnie od informacji o jego położeniu (Ungerleider, Mishkin, 1982).

## WYNIKI SYMULACJI

Model szeregowania Mareschala i Shultza (1999) przejawia wszystkie istotne cechy rozwoju szeregowania: (1) okresy stałego zachowania, odpowiadające stadiom rozwoju; (2) określoną kolejność przechodzenia przez stadia; (3) lepsze wykonanie wraz ze wzrostem różnic w wielkości elementów; (4) zmienność w strategiach wewnątrz każdego stadium, co jest zgodne z obserwacjami, że strategie wyboru przypadkowego i częściowego występują u dzieci na wszystkich poziomach rozwoju (por. Kingma, 1983); (5) stopniowe, a nie skokowe przechodzenie do wyższych stadiów, które polega na powolnym zanurzaniu się jednego stylu zachowania w drugim, co jest zgodne z badaniami Kingmy (1983), w których stwierdzono przewagę stopniowego, a nie skokowego rozwoju zdolności szeregowania. Większość znanych symbolicznych modeli szeregowania ujawnia jedynie pierwszą właściwość – okresy stałego zachowania, odpowiadające określonemu stadium.

Symulacje, w których manipulowano takimi zmiennymi zbioru uczącego, jak różnice w wielkości elementów, liczebność zbioru i stopień początkowego uporządkowania, ujawniły również, że uczenie jednego i drugiego modułu przebiega w różnym tempie. Więcej cykli uczenia wymaga zadanie „który” niż zadanie „gdzie”. Również średnia liczba „rekrutowanych” jednostek ukrytych jest większa dla tego modułu, co wskazuje, że zadanie typu „który” jest bardziej złożone niż zadanie typu „gdzie”.

Interesujący wniosek dotyczy roli jednostek ukrytych. Wiedza o tym, jakie działanie powinno zostać wykonane w odpowiedzi na określony stan wejścia, zakodowana jest w wagach połączeń między wejściem i wyjściem.

Okazało się, że włączanie nowych jednostek rzadko prowadziło do natychmiastowej zmiany strategii szeregowania. Mareschal i Shultz (1999) wyciągają stąd wniosek, że jakościowa zmiana w potencjale komputacyjnym nie musi prowadzić do jakościowej zmiany w obserwowalnym zachowaniu. Dołączenie nowej jednostki ukrytej zwiększało zazwyczaj jedynie tempo uczenia.

Wyniki symulacji pozwalają na bliższy wgląd w naturę niektórych mechanizmów odpowiedzialnych za rozwój szeregowania. Prawidłowe uporządkowanie stadiów rozwoju wynika ze stopniowego rozszerzania zdolności szeregowania, ograniczonych początkowo do krótkich serii, na coraz dłuższe serie elementów. Jest to zgodne z obserwacjami, że już dzieci w wieku 2-4 lat potrafią szeregować niewielkie zbiory, lecz nie potrafią zastosować tej zdolności do większych zbiorów (por. Kingma, 1984). Przechodzenie do następnych stadiów jest z kolei wynikiem przede wszystkim zmian zachodzących w wagach połączeń, a w mniejszym stopniu wynikiem dołączania nowych jednostek. Natomiast zmienność w strategiach wynika z tego, że strategie te wynurzają się z leżących u ich podstaw mechanizmów zmian w topologii połączeń między jednostkami.

\*

Według Piageta, kolejne stadia rozwoju są okresami zachowania odzwierciedlającymi jakościowo różne sposoby przetwarzania informacji. Przechodzenie do wyższych stadiów jest nagłe i zachodzi jednocześnie w różnych dziedzinach w stałej i uporządkowanej kolejności. Symulacje koneksjonistyczne faz rozwojowych (por. Shultz, 1991; Mareschal, Shultz, 1999) zdają się jednak wskazywać, że o ile zmiany rozwojowe mają charakter zmian jakościowych i zachodzą w uporządkowanej kolejności, o tyle nie zachodzą w sposób nagły i jednoczesny we wszystkich dziedzinach.

## KONEKSJONISTYCZNE MODELE MECHANIZMÓW ROZWOJU POZNAWCZEGO

Symulacje rozwoju procesu szeregowania wykazały, że nie zawsze do osiągnięcia końcowego stadium rozwojowego konieczny był wzrost zasobów reprezentacyjnych. Przeciwnie wyniki uzyskano w symulacjach rozwiązywania problemu „wagi” i integrowania informacji na temat prędkości, czasu i drogi, gdzie „rekrutacja” nowych jednostek była niezbędnym warunkiem osiągania kolejnych stadiów. Wyniki te sugerują, że nieciągłość behawioralnych zmian rozwojowych może być wynikiem różnych mechanizmów. Co więcej, symulacje rozwoju szeregowania, przeprowadzone przez Mareschala i Shultza, wykazały, że w przypadku niektórych sieci osiąganie kolejnych stadiów wymagało „rekrutowania” nowych jednostek, a w przypadku innych nie było takiej konieczności. Wynik ten może świadczyć o tym, że u podstaw niektórych zmian behawioralnych mogą leżeć różne ścieżki rozwoju i niektóre dzieci mogą osiągać kolejne stadia rozwoju w wyniku jakościowych zmian w potencjale komputacyjnym (wzrost zasobów reprezentacyjnych), inne zaś mogą osiągać te same stadia w wyniku jedynie ilościowych zmian w potencjale komputacyjnym (zmiany parametrów połączeń). Dalsze badania powinny dać odpowiedź na pytanie, jakie cechy struktury danych wejściowych i wcześniejszych faz uczenia wpływają na prawdopodobieństwo rozwoju według jednej lub drugiej ścieżki.

W dyskusji na temat natury zmian, które leżą u podstaw rozwoju, znane są dwa stanowiska. Zwolennicy podejścia Piagetowskiego uważają, że rozwojowi towarzyszą gwałtowne zmiany w potencjale komputacyjnym (por. Case, 1985). Według nich, rozwój polega na przechodzeniu przez kolejne stadia jakościowo różnych sposobów przetwarzania informacji. Przeciwnicy takiego podejścia uważają, że w trakcie rozwoju nie zachodzą jakościowe, lecz jedynie ilościowe zmiany w sposobach przetwarzania informacji. Rozwój wiąże się z poszerzeniem wiedzy, a jej restrukturyzacja wynika z ciągłych, małych i ilościowych zmian (por. Chi, 1990).

Uczenie kaskadowo-korelacyjne pozwala na modelowanie obu rodzajów zmian rozwojowych. Przechodzenie do kolejnych stadiów może wiązać się ze zmianami w wagach połączeń albo z poszerzeniem zasobów reprezentacyjnych sieci przez dołączanie nowych jednostek. Z symulacji przeprowadzonych przez Mareschala i Shultza (1999) wynika, że jakościowe zmiany w zachowaniu nie zawsze muszą odzwierciedlać jakościowe zmiany w sposobie przetwarzania informacji. Co więcej, jakościowe zmiany w przetwarzaniu nie muszą pociągać za sobą jakościowych zmian w zachowaniu. Ten ostatni przypadek ma miejsce wówczas, gdy zachowanie jest prawidłowe, ale mało efektywne. W wyniku jakościowych zmian w sposobie przetwarzania zmianie może ulec jedynie efektywność zachowania, podczas gdy ono samo pozostanie na tym samym jakościowo poziomie. Wniosek, iż nie wszystkie zjawiska rozwojowe przejawiające się w postaci jakościowych zmian w zachowaniu muszą wiązać się z jakościowymi zmianami w leżących u ich podstaw mechanizmach przetwarzania informacji, jest zgodny z analizami opartymi na założeniach teorii systemów dynamicznych. Wskazują one, iż niektóre jakościowe zmiany rozwojowe w zachowaniu mogą być wynikiem stopniowych, ilościowych zmian w mechanizmach przetwarzania informacji (van Geert, 1991; van der Maas, Hopkins, 1998; van der Maas, Molenaar, 1992).

Należy jednak zauważyć, że nie wszystkie strukturalne zmiany w architekturze systemu prowadzą do jakościowych zmian w sposobie przetwarzania informacji, jak to sugerują Mareschal i Shultz (1999). Na przykład dołączenie do sieci nowej jednostki ukrytej ma inne następstwa dla sieci, która dotąd nie miała warstwy ukrytej, a inne dla sieci, która taką warstwę już ma. W pierwszym przypadku zmiana strukturalna prowadzi do jakościowej zmiany w potencjale komputacyjnym, ponieważ dzięki warstwie ukrytej sieć staje się zdolna do rozwiązywania zadań typu XOR. W drugim przypadku zmiana strukturalna prowadzi do ilościowej jedynie zmiany w potencjale komputacyjnym, niekoniecznie korzystnej z punktu widzenia zadania, które sieć ma realizować. Na przykład z badań nad różnymi architekturami sieci wynika, że zbyt duża w stosunku do wielkości wektora wejściowego liczba jednostek w warstwie ukrytej nie sprzyja zdolności uogólniania.

Podsumowując, nie można jakościowych zmian w sposobie przetwarzania informacji traktować jako wyniku jedynie strukturalnych zmian w architekturze sieci, a ilościowych zmian w sposobie przetwarzania informacji jako wyniku jedynie ilościowych zmian w parametrach (wagach) połączeń. Stąd też propozycję Shultza i Schmidta (1991), aby ilościową fazę redukcji błędów i zmiany wag traktować jako odpowiednik Piagetowskiego procesu asymilacji, a fazę zmian jakościowych jako odpowiednik procesu akomodacji, należy traktować jako nadmiernie uproszczoną.

Pytanie, czy rozwój ma charakter ciągły, czy skokowy znajduje w koneksjonizmie dwojaką odpowiedź. W stopniu, w jakim u podstaw behawioralnych zmian rozwojowych leżą mechanizmy uczenia się związane jedynie z modyfikacją wag połączeń (a nie dołączaniem lub odrzucaniem jednostek) rozwój, analizowany na poziomie mechanizmów, ma charakter ciągły, ponieważ wagi połączeń zmieniają się stopniowo<sup>2</sup>. Natomiast w sieciach o zmiennej architekturze, takich jak sieci kaskadowej korelacji, zmiany mają charakter zarówno ciągły jak i nieciągły, ilościowy, jak i jakościowy. Dalsze badania powinny dać odpowiedź na pytanie, w jaki

<sup>2</sup>Pomijamy tu różnicę między przyrostowym i kumulacyjnym uaktualnianiem wag. W drugim przypadku zmiana wartości wag następuje dopiero po zakończeniu pełnego cyklu uczenia (przedstawieniu wszystkich przykładów uczących) i może mieć charakter skokowy, prowadzący do nagłej zmiany w zachowaniu.

ADAM PUTKO

sposób interakcja między wymaganiami zadania a ograniczeniami w mechanizmach uczenia może prowadzić do pojawiania się określonych zmian rozwojowych.

## BIBLIOGRAFIA

- Bates, E. A., Elman, J. L. (1993). Self-organisation and cognitive change. [W:] M. H. Johnson (red.), *Brain development and cognition* (s. 560-610). Oxford: Blackwell.
- Baylor, G. W., Gascon, J., Lemoyne, G., Pothier, N. (1973). An information-processing model of some seriation tasks. *Canadian Psychologist*, 14, 157-196.
- Brown, R. (1973). *A first language: The early stages*. London: Allien & Unwin.
- Buckingham, D., Shultz, T. R. (1994). A connectionist model of the development of velocity, time, and distance concepts. *Proceedings of the 16th annual conference of the cognitive science society* (s. 245-289). Hillsdale, NJ: Erlbaum.
- Bybee, J., Slobin, D. I. (1982). Rules and schemata in the development and use of English past tense. *Language*, 58, 265-289.
- Case, R. (1985). *Intellectual development: Birth to adulthood*. New York: Academic Press.
- Changeux, J., Dehaene, S. (1989). Neural models of cognitive functions. *Cognition*, 33, 63-109.
- Chi, M. T. H. (1990). Knowledge structures and memory. W: R. S. Siegler (red.), *Children's thinking: What develops?* (s. 135-162). Hillsdale, NJ: Erlbaum.
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of control. Signals, and Systems*, 2, 304-314.
- Dawson, G., Fischer, K. W. (red.) (1994). *Human behaviour and the developing brain*. New York: Guilford.
- Elman, J. L., Bates, E. A., Karmiloff-Smith, A., Johnson, M. H., Parisi, D. Plunkett, K. (1996). *Rethinking innateness: Connectionism in a developmental framework*. Cambridge, MA: MIT Press.
- Fahlman, S. E., Lebiere, C. (1990). The cascade-correlation learning architecture. W: D. S. Touretzky (red.), *Advances in neural information processing systems 2* (s. 181-220). San Mateo, CA: Morgan Kaufmann.
- Fodor, J. (1981). *Representations*. Cambridge, MA: MIT Press.
- Hanson, S. J. (1990). Meiosis networks. W: D. S. Touretzky (red.), *Advances in neural-information processing systems 2* (s. 533-542). San Mateo, CA: Morgan Kaufmann.
- Inhelder, B., Piaget, J. (1969). *The early growth of logic in the child*. New York: Norton Library.
- Jacobs, R. A., Jordan, M. I., Barto, A. G. (1991). Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. *Cognitive Science*, 15, 219-250.
- Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition*, 23, 95-147.
- Karmiloff-Smith, A. (1992). *Beyond modularity*. Cambridge: Cambridge Univ. Press.
- Kingma, J. (1983). The development of seriation, conservation, and multiple classification: A longitudinal study. *Genetic Psychology Monographs*, 108, 43-67.
- Kingma, J. (1984). The influence of task variations in seriation research: Adding irrelevant cues to the stimulus materials. *Journal of Genetic Psychology*, 144, 241-253.
- Kuczaj, S. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, 16, 589-600.
- MacWhinney, B. (1991). *The CHILDES project: Tools for analyzing talk*. Hillsdale, NY: Erlbaum.
- Marchman, V. (1984). Learning not to overgeneralize. *Papers and Reports on Child Language Development*, 24, 69-74.
- Marchman, V., Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, 21, 2, 339-366.
- Marcus, G. F., Ullman, M., Pinker, S., Hollander, M., Rosen, T. J., Xu, F. (1992). Overregularization in language acquisition. W: *Monographs of the society for research in child development* (s. 112-148). Oxford: Blackwell.
- Mareschal, D., Shultz, T. R. (1996). Generative connectionist algorithms and constructivist cognitive development. *Cognitive Development*, 11, 571-603.
- Mareschal, D., Shultz, T. R. (1999). Development of children's seriation: A connectionist approach. *Connection Science*, 11, 149-186.
- McClelland, J. L. (1989). Parallel distributed processing: Implications for cognition and development. W: R. G. M. Morris (red.), *Parallel distributed processing: Implications for psychology and neurobiology* (s. 35-72). Oxford: Oxford Univ. Press.
- McClelland, J. L. (1995). A connectionist perspective on knowledge and development. W: T. J. Simon, G. S. Halford (red.), *Developing cognitive competence: New approaches to process modelling* (s. 170-204). Hillsdale, NJ: Erlbaum.
- Piaget, J. (1965). *The child's conception of number*. New York: Norton.
- Piaget, J. (1972). *Strukturalizm*. Warszawa: Wiedza Powszechna.
- Piaget, J., Inhelder, B. (1973). *Memory and intelligence*. London: Routledge & Kegan Paul.
- Pinker, S., Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-193.
- Plunkett, K., Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition*, 38, 1-60.
- Plunkett, K., Marchman, V., Knudsen, S. L. (1991). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. W: D. S. Touretzky, J. L. Elman, T. J. Sejnowski, G. E. Hinton (red.), *Proceedings of the 1990 connectionist models summer school* (s. 201-219). San Mateo, CA: Morgan Kaufmann.
- Plunkett, K., Sinha, C. (1992). Connectionism and developmental theory. *British Journal of Developmental Psychology*, 10,



## KONEKSJONISTYCZNE MODELE MECHANIZMÓW ROZWOJU POZNAWCZEGO

209-254.

Quartz, S. R. (1993). Neural networks, nativism, and the plausibility of constructivism. *Cognition*, 48, 223-242.

Quartz, S. R., Sejnowski, T. J. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavioural and Brain Sciences*, 20, 537-596.

Quinn, P. C., Johnson, M. H. (1997). The emergence of perceptual category representations in young infants. *Journal of Experimental Child Psychology*, 66, 236-263.

Rumelhart, D. E., McClelland, J. L. (1986). On learning the past tense of English verbs. W: J. L. McClelland, D. E. Rumelhart, PDP Research Group (red.), *Parallel distributed processing: Explorations in the microstructure of cognition* (vol. 1, s. 318-362). Cambridge, MA: MIT Press.

Rumelhart, D. E., McClelland, J. L. (1987). Learning the past tenses of English verbs: Implicit rules or parallel distributed processing. W: B MacWhinney (red.), *Mechanisms of language acquisition* (s. 210-248). Hillsdale, NJ: Erlbaum.

Shultz, T. R. (1991). Simulating stages of human cognitive development with connectionist models. W: L. Birnbaum, G. Collins (red.), *Machine learning: Proceedings of the 8th international workshop* (s. 305-340). San Mateo, CA: Morgan Kaufmann.

Shultz, T. R., Schmidt, W. C. (1991). A cascade-correlation model of balance scale phenomena. W: *Proceedings of the 13th annual conference of the cognitive science society* (s. 635-640). Hillsdale, NJ: Erlbaum.

Shultz, T. R., Schmidt, W. C., Buckingham, D., Mareschal, D. (1995). Modeling cognitive development with a generative connectionist algorithm. W: G. Halford, T. Simon (red.), *Developing cognitive competence: A new approaches to process modelling* (s. 205-262). Hillsdale, NJ: Erlbaum.

Siegler, R. S. (1981). Developmental sequences within and between concepts. *Monographs in the Society for Research in Child Development*, 46, 1-74.

Siegler, R. S. (1995). How does change occur: A microgenetic study of number conversation. *Cognitive Psychology*, 28, 225-273.

Siegler, R. S. (1996). *Emerging minds*. Oxford: Oxford Univ. Press.

Thelen, E., Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Ungerleider, L. G., Mishkin, M. (1982). Two cortical visual systems. W: D. J. Ingle, M. A. Goodale, R. J. W. Mansfield (red.), *Analysis of visual behaviour* (s. 549-586). Cambridge, MA: MIT Press.

Van der Maas, H. L. J., Hopkins, B. (1998). Developmental transitions: So what's new? *The British Journal of Developmental Psychology*, 16, 1-10.

Van der Maas, H. L. J., Molenaar, P. C. M. (1992). Stagemwise cognitive development: An application of catastrophe theory. *Psychological Review*, 99, 395-417.

Van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, 98, 3-53.

Van Geert, P. (1994). *Dynamic systems of development: Change between complexity and chaos*. London: Harvester Wheatsheaf.

Young, R. (1976). *Seriation by children: An artificial intelligence analysis of a Piagetian task*. Basel: Birkhauser.