**Îáîáùåííàÿ ãðàììàòèêà íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ**.

Îáîáùåííàÿ ãðàììàòèêà íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ (ÎÃÍÑ) â íàñòîÿùåå âðåìÿ øèðîêî èñïîëüçóåòñÿ â ïðèêëàäíîé ëèíãâèñòèêå. Åñëè ðàññóæäàòü ìàòåìàòè÷åñêè, ÎÃÍÑ - ýòî âàðèàíò êîíòåêñòíî-ñâîáîäíîé ãðàììàòèêè íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ. Èñòîðè÷åñêè ÎÃÍÑ ðàçâèâàëàñü èç ãðàììàòèêè Ìîíòåãþ (Montague). Ãðàììàòèêîé íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ ïðåíåáðåãàëè â òå÷åíèå äâóõ äåñÿòèëåòèé. Çàòåì îêàçàëîñü, ÷òî âñå äîâîäû ïðîòèâ åå îïèñàòåëüíîé íåàäåêâàòíîñòè äëÿ åñòåñòâåííûõ ÿçûêîâ îêàçàëèñü íåâåðíûìè èëè îñíîâàííûìè íà ëîæíûõ ïðåäïîñûëêàõ. Äàííàÿ ãðàììàòèêà î÷åíü õîðîøî âîñïðèíèìàåòñÿ ïðèêëàäíûìè ëèíãâèñòàìè è ïðîãðàììèñòàìè. Â Îáîáùåííîé ãðàììàòèêå ÍÑ êîíòåêñòíî-ñâîáîäíàÿ ãðàììàòèêà ÍÑ âûðàæàåòñÿ íå â ÿâíîì âèäå, à ïîñðåäñòâîì ðàçëè÷íûõ ïðèåìîâ, êîòîðûå ïîçâîëÿþò îïðåäåëèòü ëèíãâèñòè÷åñêè çíà÷èìûå îáîáùåíèÿ è ïðåäëàãàþò ïðàâèëà óìíîæåíèÿ, áîëåå êîìïàêòíûå, ÷åì ïðîñòîå ïåðå÷èñëåíèå ïðàâèë.

**Òåîðåòè÷åñêîå ïðåäñòàâëåíèå**

ÎÃÍÑ îïðåäåëÿåò ñèíòàêñè÷åñêèå êàòåãîðèè êàê íàáîð ñèíòàêñè÷åñêèõ ïðèçíàêîâ. Ñèíòàêñ÷åñêèé ïðèçíàê - ýòî îïðåäåëåíèå, ñîñòîÿùåå èç äâóõ ñèìâîëîâ, îïðåäåëÿþùèõ ñàì ïðèçíàê (íàïðèìåð, ïàäåæ) è åãî çíà÷åíèå (ðîäèòåëüíûé), ëèáî ýòî ìîæåò áûòü ñèíòàêñè÷åñêàÿ êàòåãîðèÿ (ò.å. ïðèçíàêè ìîãóò ïðèíèìàòü êàòåãîðèè â êà÷åñòâå èõ çíà÷åíèé).Ñèíòàêñè÷åñêàÿ êàòåãîðèÿ, òàêèì îáðàçîì, ÿâëÿåòñÿ ÷àñòè÷íî ôóíêöèåé ïåðåõîäà îò ïðèçíàêà ê åãî çíà÷åíèþ. Âíóòðåííåå ñîñòàâëåíèå êàòåãîðèé îãðàíè÷èâàåòñÿ âîçìîæíûìè êîìáèíàöèÿìè ïðèçíàêîâ (Áóëåâû óñëîâèÿ). ñèíòàêñè÷åñêèå ñòðóêòóðû - ýòî äåðåâüÿ íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ, êîòîðûå îðãàíèçóþòñÿ ïî âûøåóïîìÿíóòîìó ïðèíöèïó. Ïîñòðîåíèå äåðåâà îñóùåñòâëÿåòñÿ ïî îïðåäåëåííûì ïðàâèëàì: ïðàâèëà çàâèñèìîñòè, ïðàâèëà ëèíåéíîãî ïðåäøåñòâîâàíèÿ, ïðèíöèïû èëëþñòðàöèè ïðèçíàêîâ è ïðàâèëà îïðåäåëåíèÿ îñíîâíûõ ïðèçíàêîâ.

Ïðàâèëà çàâèñèìîñòè ïîçâîëÿþò îäíîìó ñëîâó çàâèñåòü îò äðóãîãî. Íàïðèìåð, S -> NP,VP. Ïðàâèëà çàâèñèìîñòè ëèáî ïåðå÷èñëåíû, ëèáî âûâîäÿòñÿ èç äðóãèõ ìåòàïðàâèëàìè. Ïðàâèëà ëèíåéíîãî ïðåäøåñòâîâàíèÿ îïðåäåëÿþò ïîðÿäîê ïîÿâëåíèÿ êàòåãîðèé. ïðèíöèïû èëëþñòðàöèè ïðèçíàêîâ îñíîâàíû íà äóáëèðîâàíèè ïðèçíàêîâ (ò.å. ñîãëàñîâàíèå ïðèëàãàòåëüíîãî è ñóùåñòâèòåëüíîãî, ãëàãîëà è ñóùåñòâèòåëüíîãî). ïðàâèëà îïðåäåëåíèÿ îñíîâíûõ ïðèçíàêîâ ðàáîòàþò òîãäà, êîãäà êàòåãîðèÿ ìîæåò âñòðåòèòüñÿ, òîëüêî åñëè â ïðåäëîæåíèè íå âñòðåòèëàñü îòðèöàþùàÿ åå êàòåãîðèÿ.

**Ïðèìåíåíèå â ïðèêëàäíûõ öåëÿõ**

Ðàáîòà ãðàììàòèêè íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ ïîêðûâàåò òàêèå êàòåãîðèè, êàê ñèñòåìó âñïîìîãàòåëüíûõ ãëàãîëîâ àíãëèéñêîãî ÿçûêà, ñîãëàñîâàíèå, âîïðîñû, îòíîñèòåëüíûå ïðåäëîæåíèÿ, ïàññèâíûå êîíñòðóêöèè, ñî÷åòàíèÿ ñ ñóùåñòâèòåëüíûì è ò.ä. Ýòà ãðàììàòèêà áûëà ðàçðàáîòàíà íå òîëüêî äëÿ àíãëèéñêîãî ÿçûêà, íî òàêæå äëÿ àðàáñêîãî, êàòàëàíñêîãî, êèòàéñêîãî, ãîëëàíäñêîãî, ôðàíöóçñêîãî, íåìåöêîãî, ãðå÷åñêîãî, õèíäè, èðëàíäñêîãî, ÿïîíñêîãî, êîðåéñêîãî, ëàòèíñêîãî, ïîëüñêîãî, èñïàíñêîãî, øâåäñêîãî è óýëüññêîãî.

Ìíîãèå ñîâðåìåííûå ñèñòåìû èñïîëüçóþò ãðàììàòèêó íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ. Ñàìûé áîëüøîé êîììåð÷åñêèé ïðîåêò ñ èñïîëüçîâàíèåì àíàëèçàòîðà íà îñíîâå ãðàììàòèêè ÍÑ ïðèíàäëåæèò ôèðìå Hewlett Packard è íàçûâàåòñÿ “HPSG”.

**Äåðåâî íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ**

Äåðåâüÿ ÍÑ ïðåäñòàâëÿþò ñòðóêòóðíîå îïèñàíèå äëÿ ïðåäëîæåíèé. Â äàëüíåéøåì ñòðóêòóðíîå îïèñàíèå ïðåäëîæåíèÿ ìîæåò èñïîëüçîâàòüñÿ â ñèñòåìàõ, àíàëèçèðóþùèõ èëè ñèíòåçèðóþùèõ åñòåñòâåííûé ÿçûê.

Ïðèìåð äåðåâà ÍÑ äëÿ ïðåäëîæåíèÿ:

John wanted to publish the paper.

S

NP VP

VP

NPR V P V NP

DET N

John wanted to publish the paper

Ñîîòâåòñòâóþùàÿ ñòðóêòóðà ñêîáîê:

[S[NP[NPRJohn]][VP[Vwanted][VP[Pto][Vpublish][NP[DETthe][Npaper]]]]]

Äåðåâüÿ ÍÑ è ñòðóêòóðû ñî ñêîáêàìè ÿâëÿþòñÿ òåìè ñòðóêòóðàìè, êîòîðûå ïîðîæäàþòñÿ àíàëèçàòîðîì ñèñòåì, îáðàáàòûâàþùèõ åñòåñòâåííûé ÿçûê.

**Ãðàììàòèêè íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ**

Ãðàììàòèêè ÍÑ ñîñòîÿò èç íàáîðà íåêîíå÷íûõ ñèìâîëîâ (íàïðèìåð, òàêèå êàòåãîðèè, êàê N, V, DET, P, NP, S è ò.ä.), íàáîðà êîíå÷íûõ ñèìâîëîâ (ëåêñè÷åñêèå åäèíèöû òèïà “Äæîí”, “êóïèòü”, “ãàçåòà” è ò.ä.) è íàáîð ïðàâèë, êîòîðûå ïðåîáðàçóþò íåêîíå÷íûå ñèìâîëû â öåïî÷êó êîíåíûõ è íåêîíå÷íûõ ñèìâîëîâ. Åñëè äàííûå ïðàâèëà íå çàâèñÿò îò êîíòåêñòà íåêîíå÷íîãî ñèìâîëà, òî òàêàÿ ãðàììàòèêà íàçûâàåòñÿ êîíòåêñòíî-ñâîáîäíîé; â ïðîòèâíîì ñëó÷àå ãðàììàòèêà íàçûâàåòñÿ êîíòåêñòíî-çàâèñèìîé.

**Ãðàììàòèêà ÍÑ â Òðàíñôîðìàöèîííîé ãðàììàòèêå**. Ãðàììàòèêà ÍÑ èãðàåò î÷åíü áîëüøóþ ðîëü â òðàíñôîðìàöèîííîé ãðàììàòèêå. Îñíîâíàÿ èäåÿ ÒÃ çàêëþ÷àåòñÿ â òîì, ÷òî äëÿ ñòðóêòóðíîãî îïèñàíèÿ ñóùåñòâóåò áàçîâûé êîìïîíåíò è, ñîîòâåòñòâåííî, áàçîâûå ïðàâèëà. Âñå îñòàëüíûå ïðàâèëà âûâîäÿòñÿ èç áàçîâûõ ïóòåì òðàíñôîðìàöèîííûõ ïðàâèë. Äåðåâüÿ, ïîðîæäåííûå òðàíñôîðìàöèîííûìè ïðàâèëàìè, òàêæå ÿâëÿþòñÿ äåðåâüÿìè íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ. Äàííîå ïðåäñòàâëåíèå ÒÃ ÿâëÿåòñÿ ñèëüíî óïðîùåííûì, íî äëÿ äàííîé ðàáîòû åãî âïîëíå äîñòàòî÷íî.

**Âîçðîæäåíèå ãðàììàòèêè ÍÑ è äåðåâüåâ ÍÑ**

Íà÷èíàÿ ïðèìåðíî ñ 1975 ãîäà â ãðàììàòèêó ÍÑ áûëè âíåñåíû íåêîòîðûå äîïîëíåíèÿ è îáíîâëåíèÿ. Íå âûõîäÿ çà ðàìêè êîíòåêñòíî-ñâîáîäíîé ãðàììàòèêè, îáðàçîâàëàñü Îáîáùåííàÿ ãðàììàòèêà íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ (ÎÃÍÑ). Ãðàììàòèêà Ñëîæåíèÿ Äåðåâüåâ (ÃÑÄ) - Tree-Adjoining Grammar - îáðàçîâàëàñü íà îñíîâå ïîñòðîåíèÿ äåðåâüåâ áåç èñïîëüçîâàíèÿ ïðàâèë ïåðåõîäà. Ðàññìîòðèì ýòè äâå ãðàììàòèêè.

**Îáîáùåííàÿ ãðàììàòèêà íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ (ÎÃÍÑ)**. Êðîìå âñåõ âûøåóïîìÿíóòûõ ñâîéñòâ ãðàììàòèê ÍÑ Ãàçäàð (Gazdar) ââåë â ÎÃÍÑ äâà íîâûõ ïîíÿòèÿ: “ïóñòûå” êàòåãîðèè (ýëèïñèñ?) è ñâÿçàííûå ïðàâèëà, à òàêæå ìåòàïðàâèëà äëÿ âûâîäà îäíèõ ïðàâèë èç äðóãèõ. Ýëèïñèñ è ñâÿçàííûå ïðàâèëà íå óâåëè÷èâàþò ïîðîæäàþùóþ ñïîñîáíîñòü çà ïðåäåëû êîíòåêòñíî-ñâîáîäíîé ãðàììàòèêè. Ìåòàïðàâèëà, åñëè îíè íå îãðàíè÷åíû êàêèìè-ëèáî óñëîâèÿìè, óâåëè÷èâàþò ïîðîæäàþùóþ ñïîñîáíîñòü çà ïðåäåëû êîíòåêñòíî-ñâîáîäíîé ãðàììàòèêè, ïîòîìó ÷òî ìåòàïðàâèëî ìîæåò ïîðîæäàòü áåñêîíå÷íûé íàáîð êîíòåêñòíî-ñâîáîäíûõ ïðàâèë, êîòîðûå â ñâîþ î÷åðåäü ïîðîæäàþò êîíòåêñòíî-çàâèñèìûé ÿçûê. Ìåòàïðàâèëà â ðåàëüíûõ ãðàììàòèêàõ, îïèñàííûõ ïî ïðàâèëàì îáîáùåííîé ãðàììàòèêè ÍÑ, äîñòàòî÷íî îãðàíè÷åíû è ïîýòîìó íå óâåëè÷èâàþò ïîðîæäàþùóþ ñïîñîáíîñòü.

Ýëèïñèñ è ñâÿçàííûå ïðàâèëà (íóæíû) ïðåäñòàâëåíû äëÿ áàçîâîãî ïîðîæäåíèÿ “íåñâÿçàííûõ” çàâèñèìîñòåé, êîòîðûå äîëæíû áûòü ó÷òåíû â ïðåäëîæåíèè.

Ïðèìåð ñâÿçàííîãî ïðàâèëà - ýòî ïðàâèëî, êîòîðîå ïðåäñòàâëÿåò ýëèïñèñ äëÿ òîïèêàëèçàöèè, ò.å. íàõîæäåíèÿ òîïèêà â ïðåäëîæåíèè.

Ïîíÿòèå ýëèïñèñà íå íîâîå ïîíÿòèå. Åãî òàêæå óïîìèíàëè Harris, Kuno è Sager. Â êîíòåêñòíî-ñâîáîäíîé ãðàììàòèêå Kuno ñóùåñòâóþò èìåíà óçëîâ ñ ñîîòâåòñòâóþùèìè îïèñàíèÿìè, êîòîðûå îòðàæàþò ïðîïóùåííóþ ñîñòàâëÿþùóþ.

Sager ñîçäàë î÷åíü ïîëíûé àíàëèçàòîð. Åãî èäåè ïåðåïëåòàáòñÿ ñ èäåÿìè Ãàçäàðà. Íî Ãàçäàð áûë ïåðâûé, êòî âíåäðèë ïîíÿòèå ýëèïñèñà è ñâÿçàííûõ ïðàâèë â ôîðìàëüíóþ ðàìêó ñèíòàêñè÷åñêîé òåîðèè è èñïîëüçîâàë ýòî ïîíÿòèå ñèñòåìàòè÷åñêè äëÿ îáúÿñíåíèÿ ñòðóêòóðû äåðåâüåâ.

**Ãðàììàòèêà Ñëîæåíèÿ Äåðåâüåâ (ÃÑÄ).** Â îáîáùåííóþ ãðàììàòèêó íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ áûëè âíåñåíû íåêîòîðûå èçìåíåíèÿ, à èìåííî: ýòè èçìåíåíèÿ ïîçâîëÿþò ñîçäàâàòü ñòðóêòóðíûå îïèñàíèÿ ïî ïðèíöèïó òðàíñôîðìàöèîííîé ãðàììàòèêè, íî áåç èñïîëüçîâàíèÿ òðàíñôîðìàöèîííûõ ïðàâèë. Ãðàììàòèêà Ñëîæåíèÿ Äåðåâüåâ çàèìñòâîâàëà ó òðàíñôîðìàöèîííîé ãðàììàòèêè òåõíîëîãèþ ïîñòðîåíèÿ äåðåâüåâ íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ ïî îïðåäåëåííûì ïðàâèëàì ïîñòðîåíèÿ äåðåâüåâ.

ÃÑÄ ñîñòîèò èç íàáîðà “íà÷àëüíûõ” äåðåâüåâ è íàáîðà äîïîëíèòåëüíûõ äåðåâüåâ, à òàêæå îïåðàöèè Ñëîæåíèÿ. Íà÷àëüíûå è äîïîëíèòåëüíûå äåðåâüÿ âìåñòå íàçûâàþòñÿ áàçîâûìè äåðåâüÿìè.

Äåðåâî À íàçûâàåòñÿ “íà÷àëüíûì äåðåâîì”, åñëè îíî èìååò ôîðìó:

S

A =

Êîíå÷íûå ñèìâîëû

Êîðíåâîé óçåë îòìå÷åí ñèìâîëîì S, ïîãðàíè÷íûå óçëû ÿâëÿþòñÿ êîíå÷íûìè ñèìâîëàìè. Âíóòðåííèå óçëû - íåêîíå÷íûå ñèìâîëû.

Äåðåâî Â íàçûâàåòñÿ “äîïîëíèòåëüíûì”, åñëè îíî èìååò ôîðìó:

Õ

Â =

Õ

Êîíå÷íûå Êîíå÷íûå

Îïåðàöèÿ Ñëîæåíèÿ ñîñòîèò â òîì, ÷òî “íà÷àëüíîå” äåðåâî êàê áû íàêëàäûâàåòñÿ íà “äîïîëíèòåëüíîå” äåðåâî ïî êîðíåâîìó óçëó.

ÃÑÄ áîëåå ìîùíàÿ, ÷åì êîíòåêñòíî-ñâîáîäíàÿ ãðàììàòèêà. Ñóùåñòâóþò ÿçûêè, êîòîðûå ìîãóò áûòü ïîðîæäåíû ÃÑÄ, íî íå ÊÑÃ (íàïðèìåð, êîíòåêñòíî-çàâèñèìûå ÿçûêè).

ÃÑÄ èëëþñòðèðóåò, êàê äåðåâüÿ íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ ìîãóò áûòü ïîëó÷åíû èç áàçîâûõ äåðåâüåâ. Â îòëè÷èå îò òðàíñôîðìàöèîííîé ãðàììàòèêè ýòî äîñòèãàåòñÿ íå ñ ïîìîùüþ òðàíñôîðìàöèîííûõ ïðàâèë, à ñ ïîìîùüþ îïåðàöèè Ñëîæåíèÿ. ÃÑÄ ìîæåò ïîðîæäàòü äåðåâüÿ ñ ïåðåñåêàþùèìèñÿ çàâèñèìîñòÿìè.

**Ðåçþìå**

Äåðåâüÿ íåïîñðåäñòâåííûõ ñîñòàâëÿþùèõ îáåñïå÷èâàþò âîçìîæíîñòü ñòðóêòóðíîãî îïèñàíèÿ ïðåäëîæåíèé. Äåðåâüÿ ÍÑ ìîãóò áûòü ïîðîæäåíû ãðàììàòèêàìè ÍÑ. Äåðåâüÿ ÍÑ ìîãóò èñïîëüçîâàòüñÿ äëÿ õàðàêòåðèñòèêè ñòðóêòóðíîãî îïèñàíèÿ ïðåäëîæåíèé, âêëþ÷àÿ òå àñïåêòû, êîòîðûå îáû÷íî õàðàêòåðèçîâàëèñü òðàíñôîðìàöèîííûìè ãðàììàòèêàìè. Ýòè õàðàêòåðèñòèêè ïðåäïîëàãàþò âíåñåíèå èçìåíåíèé â êîíòåêñòíî-ñâîáîäíóþ ãðàììàòèêó, íå óâåëè÷èâàÿ åå ïîðîæäàþùóþ ñïîñîáíîñòü, èëè ïîðîæäåíèå ïðåäëîæåíèé èç àçîâûõ äåðåâüåâ ñ ïîìîùüþ îïåðàöèè Ñëîæåíèÿ, ÷òî âñå æå íåñêîëüêî óâåëè÷èâàåò ïîðîæäàþùóþ ñïîñîáíîñòü. Ñòðóêòóðíûå îïèñàíèÿ, ñäåëàííûå ñ ïîìîùüþ äåðåâüåâ ÍÑ, èñïîëüçóþòñÿ íåïîñðåäñòâåííî èëè â íåÿâíîì âèäå â ñèñòåìàõ, îïèñûâàþùèõ åñòåñòâåííûé ÿçûê.