

ФРАГМЕНТ МАТРИЧНОГО СЛОВАРЯ РУССКОГО ЯЗЫКА

Вопрос о создании справочной базы машинного фонда русского языка тесно связан с идеей словаря матричного типа.

В порядке эксперимента была составлена пробная матрица с корневой морфемой **-вод-** (значение по глаголу *водить*, а не омонимичное, связанное с существительным *вода*; факты чередований типа *водить – вожатый* (д//ж), *водить – сопровождать* (д//жд) при этом не учитывались). В матрице зафиксированы все слова с указанным корнем, извлеченные из Обратного словаря русского языка [1] и Словообразовательного словаря А.Н. Тихонова [2].

Матрица имеет по вертикали 44 строки: отсутствие префикса (нулевое начало), 21 простая и 22 сложные приставки. Приводим полный перечень префиксов, сочетающихся с корнем **-вод-** с указанием количества слов в порядке убывания: *с-* (22), *за-* (18), *про-* (16), *пере-* (16), *произ-* (16), *раз-* (15), *под-* (14), *от-* (12), *при-* (11), *вы-* (11), *об-* (11), *по-* (11), *до-* (10), отсутствие префикса (9), *на-* (9), *воспроиз-* (9), *в-* (8), *у-* (7), *вз-* (7), *не-* (6), *из-* (6), *воз-* (6), *сопро-* (6), *пред-* (5), *низ-*(4), *обза-* (4), *сверхпро-* (4), *препро-* (4), *непро-* (3), *повы-*(3), *непроиз-* (3), *нес-* (2), *непере-* (2), *выпро-* (2), *повз-* (2), *межза-* (2), *невы-* (1), *пораз-* (1), *перепроиз-* (1), *недопроиз-* (1), *упо-* (1), *подза-* (1), *разза* (1), *сос-* (1). 21 префикс используется в глаголах, существительных и прилагательных. 6 префиксов сочетаются только с существительными. Только один префикс **произ-** сочетается со всеми представленными в матрице частями речи: глаголами, существительными, прилагательными, наречиями (*производить, производство, производственный, производительно*).

По горизонтали в матрице с корнем **-вод-** 61 столбец (именно столько различных финалей – концов слов может следовать за данным корнем): 35 субстантивных (соответствующих существительным как части речи), 20 адъективных (соответствующих прилагательным и причастиям), 4 глагольных и 2 наречных. Приводим полный перечень финалей, сочетающихся с корнем **-вод-**, с указанием количества слов в порядке убывания: *-ить* (30), *-ивший* (27), *-иться* (25), отсутствие финали (20), *-ящий* (20), *-ящийся* (18), *-имый* (15), *-чик* (13), *-ка* (12), *-ной* (12), *-ный* (11), *-чица* (7), *-ок* (6), *-ник* (6), *-имость* (5), *-ительный* (5), *-итель* (5), *-ство* (4), *-очный* (4), *-ец* (3), *-ничать* (3), *-ительница* (3), *-ительность* (2), *-ик* (2), *-ский* (2), *-ительство* (2), *-ительно* (2), *-ковый* (2), *-ырь* (2), *-ница* (2), *-иловка* (2), *-никовый* (2), *-чатый* (2), *-ской* (2), *-ность* (1), *-но* (1), *-очка* (1), *-ня* (1), *-ительствовать* (1), *-ливый* (1), *-ничество* (1), *-ческий* (1), *-истый* (1), *-ьба* (1), *-илка* (1), *-шико* (1), *-ительский* (1), *-ырка* (1), *-ы* (1), *-ины* (1), *-ила* (1), *-ина* (1), *-овый* (1), *-нический* (1), *-ничанье* (1), *-ище* (1), *-чанин* (1), *-а* (1), *-ственный* (1), *-ственник* (1), *-ственница* (1).

Изначальные размеры матрицы могут изменяться в сторону увеличения. Так, в картотеке НИЛТиПЛ зафиксировано слово *сопроводилочка*, не отмеченное словарями. Следовательно, по горизонтали размер матрицы увеличивается на одну единицу за счет финали **-иловочка**.

В матрице имеются 2684 клетки (44 x 61). Заполненными из них оказались 306 клеток (именно столько слов с корнем **-вод-** зафиксировано в указанных словарях). Отсутствие слова, пустоты также показательны. Их анализ на уровне многих матриц позволит разработать систему запретов на сочетаемость. Кроме того, отдельные пустые клетки могут быть заполнены материалом исторических словарей, специальной лексикой и терминами, новообразованиями и окказионализмами, диалектизмами. Матричный словарь позволяет прогнозировать и появление новых слов.

Распределение слов в гнезде с корнем **-вод-** по частям речи следующее: из 306 слов 130 прилагательных и причастий (они объединены на основе общих форм изменения), 114 имен существительных, 59 глаголов, 3 наречия. При анализе слов обнаруживаются рамочные конструкции: *с-... -ивший, -ящий, -ящийся, -имый, -ный, -ной, -чатый, -овый, -нический; пере-... -ивший, -ящий, -ящийся, -имый, -ный, -ной, -ческий; раз-... -ивший, -ящий, -ящийся, -ный, -ной, -очный, -чатый, -истый* (у прилагательных); *с-...отсутствие финали, -чик, -ка, -ник, -ик, -ница, -очка, -ня, -ничество, -ничанье; за-...отсутствие финали, -чик, -ка, -чица, -ец, -ик, -иловка, -ишко, -ила, -ище, -чанин; произ-...-итель, -ство, -ительность, -ность, -ственник, -ственница* (у существительных); *с-...-ить, -иться, -ничать; пред-...-ить, -ительствовать* (у глаголов). Забегая вперед, заметим, что часть выделенных рамочных конструкций "работают" и на матрицах с другими корнями, например, с **-ход-**. Сравните: *сводивший – сходящийся, сводящийся – сходящийся, сводный – сходный, сводка – сходка, сводить – сходить* и под.

Матричный словарь позволяет прогнозировать и сложные слова. Из указанных словарей были отобраны слова со сложными основами, включающие в качестве опорного компонента **-вод-**: с нулевой финалью – 152 слова (20), *-ство* – 66 (4), *-ческий* – 47 (1), *-ный* – 41 (11), *-ящий* – 14 (20), *-итель* – 13 (5), *-ительский* – 13 (1), *-ка* – 13 (12), *-ственный* – 1 (1), *-ский* – 10 (2), *-чик* – 10 (13), *-ить* – 8 (30), *-ской* – 8 (2), *-ец* – 7 (3), *-ник* – 5 (6), *-ность* – 4 (1), *-ительный* – 4 (5), *-ительство* – 3 (2), *-ительница* – 3 (2), *-иться* – 3 (25), *-ствовать* – 3, *-ница* – 3 (2), *-ня* – 3 (1), *-имость* – 2 (5), *-имый* – 2 (15), *-ительность* – 2 (2), *-ной* 2 (12), *-ительша* – 1, *-ик* – 1 (2), *-чица* – 1 (7), *-ничать* – 1 (3), *-ительствовать* – 1 (1), *-ом* – 1, *-ами* – 1, *-ящийся* – 1 (19), *-никовый* – 1 (2), *-ительно* – 1 (2). Всего 462 слова (для сравнения в скобках воспроизведено количество простых слов с той же концовкой). Причем список сложных слов постоянно пополняется. Так, указанными словарями не зафиксированы слова *био завод, асфальто завод, аммиако завод, авто водитель, авторуководство, госплемзавод, змееводство, кукурузо заводка, мото производство, муззвод, муравьеводство, пескопровод, плем производство, полноприводной, полувзводный, путеводно, рыботрубопровод, рыбо разводный, радио заводцы, тракторо заводовец, шпало завод* и др., имеющиеся в картотеке НИЛТи ПЛ.

Приняв за основу матрицу с корнем **-вод-**, можно составить матрицы с корнями **-ход-**, **-род-**, **-год-** и под. При наложении их друг на друга выявляются общие зоны (префиксы и финали, сочетающиеся со всеми корнями или частью из них), а также специфика каждого корня. Так, префиксы *с-, за-, про-, пере-, раз-, вы-, по-, на-, в-, у-* и некоторые другие сочетаются с корнями **-вод-**, **-ход-**, **-род-**. Только с **-вод-** сочетаются префиксы

обза-, пораз-, вопроиз-, препро-, сопро-, выпро-, повз-, непроиз-, перепроиз-, недопроиз-, упо, межза-, подза-, раза-, сос-, сверхпро- (всего 16 сложных префиксов). Специфику корня *-ход-* составляют префиксы *превос-, оби-, пре-, израс-, перерас, снис-, безвы-, недо-, неис-, безыс-, необ-, прив-, неза-, непод-, запри-, опри-, нена-, предвы-, поис-, по-по-, попри-, пос-, безот-, недорас-, неизрас-, подо-, поу-, бездо-, незапри-* (всего 29 префиксов). Приставки *обезна-, прина-, анти-, без-, межпо-, обна-, изу-* сочетаются с корнем *-род-*, но не сочетаются с корнями *-ход-, -вод-*. Аналогично можно проследить общность и различие по концам слов. Так, например, с *-ход-*, но не с *-вод-* сочетаются *-чивый, -чиво, -чивость, -овать, -оваться, -ование, -ованный, -ом, -яще, -имо, -овой, -бище, -атайствовать, -атайствование, -атайство, -атай, -атайница, -ень, -ствовать, -ящность, -имец, -имка, -имица, -очек, -кий, -ко, -кость, -яга, -ячий, -очность, -ильный, -ному, -ак, -илки, -ули, -ульки, -улочный, -ульный, -ульно, -ульность, -ун, -унок, -цы, -ливость, -уном, -ливо*. Причем часть из финалей характерна только для *-ход-* (*-атай, -атайница* и под.), но *-овать, -оваться, -ованный, -овой, -ильный* и другие будут у *-ход-* общими с *-род-* (*уродовать, родовой, родильный*). Сочетаются с *-вод-*, но не с *-ход-* *-ительница, -ительский, -ырь, -ковый, -ырка, -ы, -ины, -ница, -иловка, -ила, -никовый, -чатый, -овый, -нический, -ничать, -ничанье, -ище, -чанин, -ской, -а, -ственный, -ственник, -ственница*. Некоторые из перечисленных финалей будут объединять матрицу *-вод-* с *-род-* и в то же время противопоставлять их матрице с *-ход-*: *-ительница, -ительский, -ина, -никовый, -ственный, -ственник, -ственница* (*родительский, родина, родниковый, родственный* и др.).

Сопоставление матриц позволяет определить общее и индивидуальное в словообразовательных гнездах с корнями *-вод-, -ход-, -род-*. Существуют элементы (как начала, так и финали), общие для всех трех корней *-вод-, -ход-, -род-*; есть специфические только для одного корня; есть элементы, родящие один корень с другим и в то же время отличающие два первых от третьего.

Всего возможно семь подмножеств сочетаемости начал и финалей с корнями *-вод-, -ход-, -род-*:

	вод	ход	род
1	+	+	+
2	+	—	—
3	—	+	—
4	+	+	—
5	+	—	+
6	—	+	+
7	—	—	+

Если учесть, что изначально по семантике корни *-вод-* и *-ход-* более близки друг другу (идея движения), а *-род-* в этом отношении стоит несколько особняком, о можно предположить, что у корней *-вод-* и *-ход-* будет больше общих элементов, чем специфических, нежели у *-вод-* и *-род-*, у *-ход-* и *-род-*.

ЛИТЕРАТУРА

1. Обратный словарь русского языка. – М.: Сов. энцикл., 1974.
2. Тихонов А.Н. Словообразовательный словарь русского языка: В 2 т. – М.: Рус. яз., 1985.