Kompozice je slovotvorný proces, při kterém se kombinují dvě nebo více slov, kořenů nebo kmenů do jednoho nového slova. Tento proces je doložený v mnoha různých jazycích, a často hraje roli na pomezí slovotvorby a větné skladby. Perspektiva pohlížející na tento fenomén multilingválně může tím pádem být cenná pro několik různých oborů jazykovědy, specificky morfologie, syntaxe, a typologie. V této práci se zaměřujeme na češtinu, angličtinu, němčinu, nizozemštinu, ruštinu, francouzštinu a španělštinu.

Kompozita v prvé řadě modelujeme z hlediska jejich základových slov. Tuto úlohu nazýváme dělením kompozit. Krom toho se zaměřujeme i na identifikaci kompozic, to znamená jejich rozlišení od ostatních typů slov. Práce začíná tím, že splnění těchto úloh ukážeme na češtině za pomoci hlubokého učení a stringových shod. Na témže jazyce práce potom zobecní dělení kompozit na vyhledávání základových slov tím, že prezentuje nástroj *Word Formation Analyzer for Czech*. Tento nástroj krom kompozice pokrývá i derivaci, což znamená, že jsme schopni automaticky dohledat slovotvorného předka pro slova, která mají jenom jednoho předka, a nemotivovaná slova, tedy rozpoznat, že slovo žádné předky nemá. Nakonec představujeme multilingvální nástroj *PaReNT* vykonávající téže úlohy, založený na bázi hlubokého modelu o vlastní architektuře, který v sobě kombinuje grafémové a sémantické reprezentace slov.

Práce pokračuje aplikací tohoto nástroj v kombinaci s manuální anotací na český slovotvorný datový zdroj DeriNet a zároveň vydáváme verzi 2.2. Dosud byl tento datový zdroj téměř výlučně orientovaný na derivaci, ale nyní je obohacen o informace o kompozici. Zároveň diskutujeme o mnoha úvahách a rozhodnutích, která byla během této činnost učiněna.

Nakonec nabízíme přehled momentálního zpracování kompozit v rámci Universal Dependencies v pěti jazycích (angličtině, češtině, němčině, niozemštině, ruštině a latině), a navrhujeme způsob, jakým kompozita modelovat pomocí závislostních struktur a nakonec přimo vtělit takto modelovaná kompozita do syntaktických struktur v tomto datovém zdroji.