

UNIVERZITA KARLOVA V PRAZE

PEDAGOGICKÁ FAKULTA

KATEDRA PSYCHOLOGIE



Bakalářská práce

Porozumění přechodně víceznačným větám

Příprava psychologického experimentu

Autorka: **Tereza Kučerová**

Vedoucí práce: Mgr. Jiří Lukavský, Ph.D.

Praha 2010

Prohlášení

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně s využitím literatury, kterou průběžně v textu cituji a její seznam uvádím ke konci práce.

V Praze 8. 4. 2010

.....
Tereza Kučerová

Poděkování

Mé poděkování patří pánům z Psychologického ústavu AV ČR, vedoucímu mé bakalářské práce *Mgr. Jiřímu Lukavskému, Ph.D.* a konzultantovi a zároveň vedoucímu experimentu *PhDr. Filipu Smolíkovi, Ph.D.*, za cenné rady a připomínky k mé práci a metodické vedení při sběru dat.

Tereza Kučerová

Abstrakt

Práce se věnuje problematice **porozumění přechodně víceznačným větám** v češtině. Seznamuje čtenáře s problematikou čtení víceznačných vět v teoretické rovině, představuje některé základní teorie, které se snaží jejich zpracování na kognitivní úrovni vysvětlit. V praktické části bylo na experimentálním materiálu, který se skládá z konkrétního typu syntakticky víceznačných vět, ověřováno, zda je *čtení těchto vět dle předpokladu pomalejší než čtení strukturně (i významově) obdobných vět jednoznačných*. Dále jsem se snažila najít *faktory, které mohou dobu čtení víceznačných vět ovlivnit*, konkrétně jsem soustředila pozornost *na délku věty a obtížnost její srozumitelnosti*.

Z metodologického hlediska jsem po přípravné fázi, kdy bylo potřeba vytvořit samotný experimentální materiál (dva soubory 35 a 40 víceznačných vět, které se lišily typem gramatické struktury), realizovala **dotazníkové šetření**, kdy jsem oslovila dohromady 83 studentů gymnázia s dvěma verzemi dotazníků (podle typu struktury vět), a následně jsem administrovala 38 pokusným osobám **počítačový experiment** (obsahující už jen jeden typ vět), jehož základem byla metoda „čtení slova po slově vlastním tempem“ tedy **metoda self-paced reading** patřící do rodiny tzv. online metod. Získaná data (týkající se pouze jednoho typu vět) jsem užila při statistické analýze za účelem zodpovězení položených otázek. Na základě výpočtů se mi potvrdilo, že doba čtení víceznačných vět je skutečně delší než doba čtení vět jednoznačných. Pokud jde o faktory, u nichž jsem sledovala vliv na dobu čtení víceznačných vět, potvrdila se mi závislost mezi dobou čtení věty a její délkou, ovšem doba čtení věty a obtížnost porozumění větě se ukázaly jako nezávislé. V práci jsem se dále pokusila v teoretické rovině nalézt možná vysvětlení pro své závěry a zároveň jsem polemizovala o některých dalších faktorech, které mohly získaná data ovlivnit.

Abstract

The thesis is focus on **the comprehension of the locally ambiguous sentences** in the Czech language. It introduces the problem of reading ambiguous sentences at the theoretical level, and also it introduces some basic theories, which try to explain the relevant cognitive processing. In the empirical part it was inspected *whether reading the ambiguous sentences is slower than reading the structural (and semantical) analogous unambiguous sentences*. Further, I tried to identify *the factors, which can influence reading time of ambiguity sentences*, more specifically I focused on *the length of sentences and difficulties of sentences comprehension*.

Concerning the methodological phase, first the experimental material was created. It consists of two sets (35 and 40 in total) of ambiguous sentences with two different types of grammatical structures. Further, I designed and administered **the questionnaire evaluation** (to 83 grammar school students). Later, **the computer experiment** has been conducted on the sample of 38 persons, which was based on **the self-paced reading method**. I used the data obtained for statistical analyses to address the hypotheses. Among the major findings, it was confirmed that reading time of ambiguous sentences is longer than reading time of unambiguous sentences. Further, the factor of the sentence length proved to have an impact on the relation of ambiguous sentences on the reading time. However, the reading time of sentence and difficulty of sentence's comprehension proved to be uncorrelated. I try to link my empirical findings with the current level of theoretical research and discuss other factors possibly having an influence on the data obtained.

Obsah

Úvod	6
Teoretická část – vysvětlení pojmů	8
I.I Teorie porozumění větám	8
I.II Víceznačné věty, přechodně víceznačné věty (<i>garden path sentences</i>)	9
I.III <i>Online methods (self-paced reading)</i> – porozumění větám v reálném čase	14
I.III.I Vnímání textu při čtení – percepční rozsah (<i>perceptual span</i>)	19
I.IV Metoda <i>syntaktického primingu</i>	20
Praktická část – příprava a realizace experimentu	23
II.I Vymyšlení vět	25
II.I.I Věty A	25
II.I.II Věty B	27
II.II Sestavení dotazníků A a B	28
II.III Pilotní výzkum	30
II.IV Příprava samotného experimentu – počítačový program	33
II.V Administrace počítačového experimentu	33
II.VI Statistické zpracování dat a interpretace výsledků	35
Diskuze	45
Závěrečné shrnutí	50
Slovníček pojmů	53
Literatura	56
Přílohy (Seznam příloh)	57

Ú v o d

Začnu-li trochu ze široka, tematicky spadá má práce do oblasti **psycholingvistiky** – je tedy na pomezí psychologie a lingvistiky, tudíž operuji s pojmy typickými pro mluvu jazykovědců, a proto pro ujasnění uvádím stručný „slovníček“ pojmů. Z teoretického pohledu bych práci dále zařadila do oblasti **kognitivní psychologie**, která mimo jiné studuje kognitivní procesy zpracování jazyka a řeči, konkrétně porozumění gramatické struktury věty.

Konkrétně je pak má bakalářská práce (tedy hlavně její *Praktická část*) součástí experimentu realizovaného na *Psychologickém ústavu Akademie věd ČR v Praze*. Už z názvu příslušného oddělení vyplývá, že zvolená oblast zájmu je polem kognitivní psychologie. Experiment se týká porozumění víceznačným větám – konkrétně, zda je při čtení víceznačných vět možno pozorovat jev syntaktického primingu – tedy, zjednodušeně řečeno, zda přečtení jedné věty ovlivní (rychlost) čtení věty následující. Syntaktická víceznačnost, se kterou zde pracujeme, vyplývá z gramatické struktury věty, respektive z tvaru (jednoho) slova, které na určitou gramatickou strukturu odkazuje. Takto víceznačné věty můžeme rozčlenit do několika skupin podle „typu“ gramatické „odchytky“, která víceznačnost způsobuje. Já osobně jsem pracovala s dvěma typy vět, ač v celém experimentu bylo použito typů více.

S ohledem na předepsaný rozsah bakalářské práce by bylo nereálné zpracovat adekvátně výsledky experimentu z hlediska potvrzení či nepotvrzení výskytu jevu syntaktického primingu. Rozsah práce by bylo nutno překročit. Ukázalo se tedy rozumné zaměřit se na některé užší, dílčí jevy, pro které jsou v rámci realizace experimentu získána data. Zaměřila jsem se tedy na rozdíly v **době čtení vět víceznačných a jim odpovídajících vět jednoznačných**, které byly též součástí experimentu (každá víceznačná věta měla v experimentu svůj jednoznačný „protiklad“, kdy se věty zpravidla lišily pouze jedním slovem a to právě tím, které vyvolávalo u jedné z vět svým tvarem přechodnou (nebo lépe – dočasnou) syntaktickou víceznačnost – tedy víceznačnost, která nastává pouze v určitém místě věty, ale ke konci věty je již vše uvedeno „na pravou míru“).

Dále jsem zjišťovala, zda existuje nějaká závislost mezi dobou čtení víceznačných vět a **hodnocením jejich srozumitelnosti**, které bylo získáno z dotazníkového šetření. Zde je hodnota srozumitelnosti považována za nezávislou proměnnou, tedy za trvalejší charakteristiku. Oproti tomu doba čtení zde figuruje jako proměnná závislá, tedy ta, která je ovlivňována manipulací některých vnějších (či vnitřních) podmínek. Jednou z takovýchto podmínek je například **délka věty** – tedy počet znaků, přičemž právě tento vliv (jako další nezávislou proměnnou) při výpočtu (statistickém zpracování dat) sleduji.

V teoretické části se snažím uvést čtenáře do problematiky studia porozumění větám, uvádím hlavní teorie, které byly v tomto kontextu vysloveny – samozřejmě se zaměřením na věty víceznačné. Jistý prostor věnuji seznámení s tzv. on-line metodami, mezi něž patří metoda self-paced reading, která byla v experimentu použita. A krátký prostor je věnován i obecné zmínce o podstatě syntaktického primingu, na který, jak už jsem zmínila, byl experiment jako celek zaměřen.

V praktické části potom začínám popisem přípravné fáze, jejíž náplní byla tvorba souboru víceznačných vět. Následuje popis sběru dat pro ověření položených hypotéz pomocí dvou metod – dotazníkové šetření a realizace počítačového experimentu pracujícího s on-line metodou self-paced reading. Práce je zakončena analýzou sebraných dat a diskuzí příslušných závěrů.

Teoretická část – vysvětlení pojmů

I.I	Teorie porozumění větám	8
I.II	Víceznačné věty, přechodně víceznačné věty (<i>garden path sentences</i>)	9
I.III	<i>Online methods (self-paced reading)</i> – porozumění větám v reálném čase	14
I.III.I	Vnímání textu při čtení – percepční rozsah (<i>perceptual span</i>)	19
I.IV	Metoda <i>syntaktického primingu</i>	20

I.I Teorie porozumění větám

Čteme-li nějaký text, neděláme to obvykle čistě bezúčelně, ale proto, abychom se z něho něco dozvěděli. Mnohým by se mohlo zdát, že to nejtěžší na celé věci je naučit se ve škole znát písmenka a spojovat je do slabik, slov a vět, no a pak už to jde vše tak nějak samo. Asi vás nepřekvapí, že opak je pravdou. Stejně jako za mnoha jinými činnostmi, i za čtením stojí složitá činnost mozku – psychologicky řečeno: kognitivní procesy.

Za porozuměním čtenému textu se skrývají dva hlavní procesy (Eysenck, 2008, s. 409) – **analýza syntaktická** neboli parsing (definováno jako „připisování gramatické struktury větě“) a **analýza sémantická** čili analýza významu věty, čemuž se blíže věnuje disciplína zvaná sémantika (resp. pragmatika soustředíme-li se jen na význam zamýšlený nikoli obecně na význam doslovný). V následujících řádcích bych se však chtěla blíže zaměřit spíše na strukturu syntaktickou. Dovolím si tedy oba procesy, alespoň v teoretické rovině, oddělit. Ostatně sami autoři se podle Eysencka (2008) na vztahu těchto dvou procesů doposud jednoznačně neshodli. V úvahu přicházejí čtyři varianty – jeden proces předchází druhému, přičemž ten dřívější ovlivňuje následující, obě analýzy probíhají zároveň, avšak paralelně vedle sebe a nebo spolu souvisejí tak těsně, že se ani nedá hovořit o dvou procesech, ale spíše o dvou podsložkách procesu jednoho. Smolík (2008, s. 1) dále zmiňuje jako jedno z hlavních témat, zabýváme-li se porozuměním jazyku, to „zda lidé reprezentují gramatickou strukturu nezávisle na konkrétních slovech, nebo zda jsou tyto reprezentace **lexikálně** specifické“.

Badatelé usilující o odhalení zákonitostí parsingu na kognitivní úrovni narážení na problém, který spočívá ve velké rychlosti průběhu příslušných procesů. Tuto „komplikaci“ se snaží obejít jednak použitím speciálních metod, o kterých se zmiňuji o několik odstavců dále, jednak tím, že přistupují k tématu oklikou a zaměřují pozornost na problémy, které mají lidé v určitých specifických situacích, které jsou nějak komplikované, např. jaké „chyby“ dělají čtenáři při čtení vět víceznačných. Vycházejí zde přitom z předpokladu, že procesy probíhající při **zpracování vět jednoznačných a víceznačných** jsou obdobné. Podobně uvádí např. Smolík (2008), že i zájem o výše zmíněné rozlišení, tedy zda se realizuje reprezentace gramatické struktury v rovině lexikální či nikoli, souvisí právě s výzkumy, které se věnují vnímání určitých typů víceznačných vět. S víceznačnými větami pracujeme i v mnou prezentovaném experimentu.

I.II Víceznačné věty, přechodně víceznačné věty (*garden path sentences*)

V případě počátečního rozčlenění víceznačných vět, které uvádí Eysenck (2008), se pohybujeme na poli „běžné“ jazykovědy. Přeskakuji zde popis rozdílu **víceznačnosti sémantické a syntaktické** za předpokladu, že rozdíl je čtenáři znám už z definice těchto dvou oblastí lingvistiky, a věnuji se rovnou větám, které jsou víceznačné syntakticky. Už samotnou víceznačnost lze rozlišit na **víceznačnost globální a lokální**. Věty, které jsou víceznačné na úrovni globální, zůstávají víceznačnými až do konce. Tedy i po dočtení věty existují minimálně dvě možné interpretace. Správný smysl se můžeme dozvědět pouze z kontextu výpovědi. Oproti tomu věty, které jsou víceznačné na úrovni lokální, jsou víceznačné jen na určitou chvíli, jen v určitém místě. Čteme-li takovou větu, dostaneme se v průběhu čtení do místa, kde se nám vyjeví víceznačnost, ale než dočteme větu do konce, vše se nám vyjasní. Jako celek je tedy věta jednoznačnou, víceznačná je jen její určitá část.

Následně se dostáváme k samotnému procesu zpracování víceznačných vět. Teorie lze rozdělit do několika skupin. Eysenck (2008) nás seznamuje s rozdělením, které vychází z výše zmíněné teze, že zpracování věty zahrnuje parsing a připisování významu. V souvislosti se zpracováním víceznačné věty pak rozděluje teorie na ty, co kladou důraz na samostatnou analýzu syntaktickou, tedy teorie sériového

zpracování věty, teorie kladoucí důraz na význam neboli teorie obsahové a teorie kombinující analýzu syntaktickou a sémantickou, tedy teorie paralelního zpracování věty. Jednotlivé typy teorií se nevyskytují pouze v čisté krystalické verzi, existují různé varianty a obměny s ohledem na to, co daný autor považuje (či považoval) za významný prvek v rámci analýzy. Z hlediska dělení teorií na sériové a paralelní se pohybujeme na škále **modularity** a **interaktivity**. Modulární sériové zpracování probíhá nejprve na úrovni lexikální, posléze probíhá parsing a až nakonec je zpracování završeno sémantickou analýzou. Oproti tomu interaktivní paralelní zpracování předpokládá realizaci pouze jednoho procesu, který je schopen naráz zkombinovat znalosti lexikální, syntaktické, sémantické a znalosti o světě.

Ke zpracování víceznačných vět bylo dále zjištěno např. i to, že preferovaná interpretace některých typů víceznačných vět je různá v různých jazycích (Smolík, 2008). Možným vysvětlením je četnost alternativních významů určité struktury v konkrétním jazyce – jedná se o tzv. **hruboornou frekvenci struktury** (Mitchell, Cuetos, Corley a Brisbaert, in Smolík, 2008), která vyzdvihuje význam frekvence syntakticky shodných vět bez ohledu na jejich lexikální skladbu.

Teorie sériového zpracování, model slepé koleje

Teorie sériového zpracování (Eysenck, 2008) říká, že si již na počátku čtení zvolíme jednu syntaktickou analýzu věty a pokud pak v průběhu či na konci čtení zjistíme, že takovéto vyložení věty není možné, vybereme si jednoduše nějakou jinou, vhodnější analýzu.

Nejvýznamnějším sériovým přístupem k syntaktické analýze věty je model slepé koleje. **Model slepé koleje** (tzv. *garden-path model*, Frazierová a Rayner, 1982, in Eysenck, 2008) získal svůj název podle toho, že víceznačná věta „svádí čtenáře na slepou kolej“ (Eysenck, 2008, s. 410). Na počátku analýzy bereme v úvahu pouze jednu a to tu nejjednodušší syntaktickou strukturu, kterou volíme bez ohledu na význam věty.

Při výběru nejjednodušší syntaktické struktury se držíme dvou principů. Prvním je tzv. **princip minimalizace (strukturních) uzlů**, druhý se nazývá **princip oddalovaného uzavření**. Strukturními uzly myslíme jmenné či slovesné fráze. Tento model se pěkně demonstruje na anglických větách, kde jsou např. předmět i počátek

předmětné věty ve shodném tvaru, na základě zmíněného principu volíme strukturu s nejmenším počtem strukturních uzlů, tedy automaticky předpokládáme, že věta bude pokračovat předmětem, nikoli že navazuje vedlejší věta předmětná. Gramatická struktura nám to bez problému umožňuje, čteme-li však dál, můžeme zjistit, že naše interpretace věty byla mylná. Obdobně v případě oddalovaného uzavření připojujeme nová slova k právě zpracovávané frázi, je-li to gramaticky přípustné.

Teorie paralelního zpracování, model omezujících faktorů

Teorie paralelního zpracování (Eysenck, 2008) staví na předpokladu, že již na počátku čtení uvažujeme několik analýz najednou a z těch nám v průběhu čtení následně vybyde ta jediná správná, která je z hlediska obsahu vyhovující.

Mezi teoriemi paralelního zpracování je tím nejvýznamnějším přístupem poměrně mladá, ale zato stále vlivnější **teorie omezujících faktorů** (MacDonaldová, Pearlmutter a Seidenberg, 1994, in Eysenck 2008). Podle této teorie má systém provádějící parsing k dispozici všechny relevantní zdroje informací či omezující faktory. Najednou je aktivováno několik soutěžících analýz, které se řadí v závislosti na síle své aktivace. Dochází zde k interakci zdrojů, která umožňuje interpretaci na každé úrovni, v průběhu čtení věty pracujeme se syntaktickými i sémantickými vodítky. Pro rozřešení víceznačnosti je zde operováno se čtyřmi pravidly. Za prvé, naše gramatické znalosti jsou omezujícími faktory pro možné interpretace věty. Za druhé, různé druhy informací asociované s jednotlivými slovy nebývají nezávislé. Za třetí, víceznačnost slova nemusí být univerzální z hlediska mluvnických kategorií, např. může být slovo víceznačné podle mluvnického času, ale už ne podle mluvnické osoby. A za čtvrté, jednotlivé interpretace, které by byly možné podle gramatických pravidel, se liší frekvencí a pravděpodobností v důsledku naší předchozí zkušenosti.

Oproti modelu slepé koleje připouští teorie omezujících faktorů **vliv kontextu** při výběru správné interpretace. Kontext aktivuje určité informace, které nám následně zpracování věty usnadňují, ale mohou ho i narušit.

Obsahové teorie zpracování – teorie kapacity

Teorie kapacity

Teorie kapacity (Just a Carpenterová, 1992, in Eysenck, 2008) je jednou z teorií, které se zabývají obecně vlivy omezujícími porozumění větám, tedy nejen těm víceznačným. Soustřeďuje se přitom konkrétně na **pracovní paměť**. Východiskem je Baddeleyho koncept, kdy je pracovní paměť vnímána jako součást centrálního vykonavatele, která se zabývá porozuměním jazyku – tedy pracovní paměť se při porozumění jazyku používá jak pro ukládání tak pro zpracování informací (Just a Carpenterová, 1992, in Eysenck, 2008, s. 416). Nároky na pracovní paměť jsou tedy poměrně vysoké. Teorie však předpokládá, že její kapacita je omezená, což má na zpracování jazyka a porozumění mu významný vliv. Například se předpokládá bezprostřední zpracování aktuálně čteného slova, místo aby bylo jen uloženo pro zpracování pozdější. I zde se, jako u většiny kognitivních jevů a schopností, předpokládá interindividuální variabilita v rozsahu kapacity a tedy i v porozumění jazyku. U jednotlivců je možné měřit konkrétní velikost kapacity testem **rozsahu čtení** (Daneman a Carpenter, 1980, in Eysenck, 2008).

Aplikace teorie kapacity na víceznačné věty, vliv významu věty

Just a Carpenterová (1992, in Eysenck, 2008) se pokusili aplikovat teorii kapacity i na rozřešení problému zpracování syntakticky víceznačných vět. Vyšli ze dvou obecných předpokladů, kdy lidé při setkání s víceznačnou větou berou v úvahu všechny možné interpretace, dokud se neobjeví rozhodující informace, nebo si zvolí jen jednu interpretaci a s tou pracují, dokud se neobjeví informace, která tuto interpretaci případně vyloučí. V praxi vyšli ze studie, kdy pokusné osoby četli jednoznačné a víceznačné věty pomocí pohyblivého okénka, zatímco experimentátor měřil dobu čtení jednotlivých slov. Z výsledků vyvodili, že osoby s velkým rozsahem čtení jsou schopné udržet několik možných interpretací najednou, zatímco osoby s malým rozsahem čtení interpretují všechny věty od počátku jako jednoznačné, tedy uvažují pouze jednu možnou interpretaci, dokud nenarazí na problematický bod.

Zároveň v témže roce realizovali experiment, který měl odpovědět na otázku, zda je počáteční parsing ovlivněn významem věty. Následně přišli s tvrzením, že osoby s velkým rozsahem čtení mají dostatečnou kapacitu pracovní paměti, aby mohli při hledání správné interpretace věty operovat i s dalšími vodítky, která vyplývají z významu jednotlivých slov.

Zhodnocení teorií

Obecně shrnuje Eysenck (2008) několik základních kritik k **modelu slepé koleje**, který je na základě řady novějších studií pomalu vyvracen. První kritika poukazuje na jeho neefektivitu, kdy by čtenáře často vedl k chybné analýze, za druhé poukazují kritici na opomíjení vlivu významu resp. kontextu a do třetice přicházejí kritici s předpokladem, že výběr počáteční struktury nezávisí jen na principech minimalizace uzlů a oddalovaného uzavření, ale zřejmě i na interpunkci (resp. prozodii v případě poslouchání nikoli čtení). Oproti tomu obhájci modelu slepé koleje (Eysenck, 2008) vyzdvihují minimální nároky na pracovní paměť a také fakt, že byl model podpořen výsledky ze sledování očních pohybů (metoda eyetrackingu).

Řada kritik se objevuje i v souvislosti s **modelem omezujících faktorů**, ty upozorňují obecně na řadu nedostatků ve výkladu zpracování syntaxe. Obdobně poukazují někteří i na **teorii kapacity**, že má sklony ignorovat různé specifické procesy účastníků se porozumění. Obhajobou modelu omezujících faktorů je zde to, že model je relativně „nový“ a je zde tedy pravděpodobné, že s dalšími výzkumy zastánci modelu mezery v teorii uspokojivě doplní (McKoonová a Ratcliff 1988, Crocker 1999, in Eysenck, 2008).

Téma kritik pěkně **shrnuje** Crocker (1999, in Eysenck, 2008, s. 413): „každý model je něčím přitažlivý a relativně dobře vysvětlující empirické nálezy, do jisté míry proto, že mnoho modelů je specifikováno a implementováno jen částečně“, s trochou nadsázky by se tedy dalo říci, že autoři interpretují výsledky svých studií tak, aby zapadly do jejich teoretického vysvětlení problematiky, případně se dívají na dílčí téma a nesnaží se vytvořit koherentní teorii, která by zahrnuje všechny oblasti dané problematiky, „sériové modely mají některé výhody, jsou konceptuálně i komputačně jednodušší (mají menší nároky na zpracování a paměť) a formulují silné a testovatelné předpovědi lidského chování.“ Mají ale

i podstatná omezení – mnohé jiné modely vysvětlují empirické nálezy lépe. Pickering (1999, in Eysenck, 2008, s. 414) závěrem **předpovídá**, že „se rýsuje takový systém, který integruje různé zdroje informace v reálném čase, aby pak volil interpretaci na základě všech těchto zdrojů.“

I.III *Online methods (self-paced reading)* – porozumění větám v reálném čase

Jak už jsem se na několika místech práce zmínila, hlavní použitou experimentální metodou je tzv. metoda **self-paced reading**. Tato metoda patří do skupiny specifických metod zvaných **on-line metody**. Tyto metody (Mitchell, 2004) vznikly z podnětu výzkumníků zabývajících se tematikou zpracování vět. Uvědomili si, že se nejedná o jednolitý proces, co by probíhal beze změny napříč celou větou. Je tedy nutno podívat se „dovnitř“ tohoto procesu, rozložit průběh zpracování věty na části. Už zde se objevují rozpory odpovídající distinkci v obecných teoriích porozumění větám. Jedna skupina zastává názor, že čtenář je schopen operovat velmi rychle s mnoha informacemi, které přispívají ke správné interpretaci věty. Druhá skupina (modulární větve) tvrdí, že čtenář sbírá informace postupně, tak jak postupuje ve větě. Už jen aby bylo možné rozhodnout, který z přístupů je správný, vynořila se potřeba nové experimentální metody, která by umožnila sledovat průběh odhalování smyslu věty. A přesně k tomu slouží on-line metody. Doposud užívané metody byly nedostačující především proto, že operace probíhající během zpracování věty trvají velmi krátkou dobu, řádově několik milisekund, bylo zde tedy riziko, že by nějaká součást jevu mohla být přehlédnuta. Sám Mitchell (2004, s. 17) k tomu říká, že „bez přístupu k on-line metodám by byli výzkumníci slepými okolostojícími spíše než plnohodnotnými svědky takovýchto aktivit“.

Historie i současnost – nejstarší, nejpoužívanější a nejnovější metody

Údajně není možné (Mitchell, 2004) sumarizovat veškerou práci, která byla realizována pomocí on-line metod, a to už pro širší tematického záběru – tyto

techniky byly užity na téměř každou otázku týkající se zpracování vět. Za dobu, kdy výzkumníci usilují o proniknutí do bezprostředních jazykových procesů užitím on-line metod, byl navíc předložen i bezpočet jejich odlišných verzí a obměn. Nabízí se způsob (Mitchell, 2004, s. 15-17) roztrždit toto velké množství pro větší přehlednost do tří skupin podle následujícího kritéria: *On-line metody se zjednodušeně „vyvíjely“ od jednoduchých a „hrubých“ po složité a sofistikované ale zato nákladné na realizaci.* Tento rozvoj koresponduje s rozvojem techniky.

Časné metody, které jsou v současné době užívány spíše jen výjimečně

Jako příklad uvádí Mitchell (2004) dvě metody, a to *the click-migration task* (pokusná osoba poslouchá z nahrávky větu, která je v určitém místě doplněna o „rušivý“ zvuk, následně má za úkol vyznačit do tištěné podoby věty místo, kde si myslí, že zvuk slyšely) a *the phoneme-monitoring task* (pokusné osoby opět poslouchaly z nahrávky větu a jejich úkolem bylo stisknout klávesu ve chvíli, kdy slyšely částečný foném). Důvodem, proč tyto a další metody upadly v zapomnění je několik, za prvé je to možná problematika v jejich interpretaci, za druhé kvůli podezření, že nálezy získané užitím některých těchto metod byly zkresleny na základě „deformace“ vycházející z metody samotné a nebo, ze zcela prostého důvodu, proto, že se výzkumníkům „dostaly do rukou“ jiné, a to často jednodušší, metody vhodné k uchopení zajímavých otázek. I přes tyto pochybnosti o použitelnosti metod, resp. o platnosti závěrů jimi získaných jsou i nadále mnohé z výsledků stále široce uznávány.

Metody známé již 20 – 30 let, které jsou ale stále aktivně používány

Zde se dostáváme k mnou použité metodě self-paced reading, kterou však podrobněji rozeberu až v následujících odstavcích. Jako další metody patřící do této skupiny jmenuje Mitchell (2004) eyetracking, který je užíván nejen pro psycholinguvistické (on-line) studie porozumění, ale i pro bezpočet dalších procesů a jevů, a užití *event-related* potenciálů (ERPs) v elektroencefalografických (EEG) studiích – výhodou této metody je, že mimo jiné poskytuje přirozený způsob studia toho, jak je materiál zpracováván, pokud je prezentován ve sluchové modalitě.

Zajímavá je zmínka (Mitchell, 2004, s. 23-24) o tendenci navracení se témat, kdy byla určitá problematika zprvu zkoumána za použití metody self-paced reading. Po několika letech se někdo k dané otázce navrátil a zopakoval výzkum s užitím metody eyetrackingu. Po další časové prodlevě se eventuálně experiment na dané téma opět opakoval, nyní již s užitím ERP techniky. Mitchell (2004) zmiňuje hned několik takových témat, kde tato posloupnost proběhla, přičemž zdůrazňuje postupné potvrzování výsledků, kdy ke stejným závěrům došli jak výzkumníci užívající jednodušší metody, tak i ti užívající moderní techniku.

Metody z posledního desetiletí, které jsou teprve podrobovány testování a hodnocení

Asi nejvlivnější taková technika je metoda užívající eyetracking, který je však nově přizpůsoben k monitorování očních pohybů v případě zobrazení nikoli tištěného textu ale vizuálních sad obsahujících jednoduché (kreslené) objekty – tato adaptace metody nám poskytuje přesnou reflexi reálného jazykového chování. Druhou takovou techniku tvoří „rodina“ neuroimaginativních metod, které v poslední době získávají stále více na oblibě (např. *positron emission tomography* (PET), *functional magnetic resonance imaging* (fMRI)). Problémem však je, že tyto metody jsou většinou příliš „pomalé“ na to, aby mohly sloužit jako on-line metody, výjimkou jsou snad jen práce užívající magnetoencephalographic (MEG) – tedy neuroimaginativní metodu, která má potenciál vyrovnat se přísným nárokům on-line výzkumů na měření času.

Self-paced reading – podstata a obměny metody

Tato relativně jednoduchá metoda (Mitchell, 2004) byla původně představena nezávisle na sobě hned několika výzkumníky a od té doby je užívána v různých variantách. Obecným základem je segmentace textu do slov, řetězců slov nebo frází – pokusné osobě je (na obrazovce počítače) v jednom okamžiku promítán vždy jen jeden segment. Experiment je spuštěn stisknutím klávesy, které vyvolá první „obraz“, jakmile osoba přečte první segment, opět stiskne klávesu a zobrazí se druhá část –

druhý segment ... a tak dále až na konec textu. Výzkumníka přitom zajímá délka časového úseku mezi „po-sobě-jdoucími“ stisky kláves – zvláště ve určitých úsecích textu.

Jak už jsem se zmínila několik řádků nazpět, existuje několik obměn metody. Jako první bych zmínila tzv. „*stationary window*“ verzi (stacionární či statické okno), kdy každý následující segment se zobrazí v té samé pozici na monitoru, jako byl ten předchozí, tedy – segmenty se postupně „přepisují“ jeden přes druhý, samozřejmě za předpokladu, že ten předchozí při zobrazení následujícího zmizí – důležité zde je tedy to, že všechny segmenty (všechna slova), která si postupně zobrazujeme, se nám ukazují na stále stejném místě. Druhá je tzv. „*moving-window*“ verze (tzv. pohyblivé okno). Zde se pozice postupně zobrazovaných segmentů pohybuje zleva doprava – každý segment textu zabírá na monitoru takovou pozici, jakou by zabíral reálně v ucelené větě. V rámci této druhé verze jsou možné dva podtypy zobrazování segmentů textu. Je-li zobrazení kumulativní, pak předchozí segmenty zůstávají na monitoru, zatímco postupně přibývají další. Je-li zobrazení nonkumulativní, pak starý segment zmizí v okamžiku, kdy se má objevit segment následující, vždy tedy vidíme jen jeden – ten aktuálně čtený segment. V obou případech se užívá většinou pomlček (či nějakých ekvivalentních značek) k označení pozice všech segmentů (slov) textu – tedy toho „zbytku“ textu, který není zrovna zobrazován. Kumulativní verze s sebou přináší specifický problém, jelikož mnohé pokusné osoby mají tendenci tisknout klávesu několikrát rychle po sobě, dokud není na monitoru zobrazena celá věta, a teprve pak předložený materiál přečte jako celek – v tomto případě však experiment postrádá svůj smysl, záměrem experimentátora je zde měřit dobu mezi jednotlivými stisky, kdy časová latence odpovídá době čtení segmentu, to však pokusná osoba maří. Oproti tomu nonkumulativní verze tento problém postrádá.

Výhody a kritiky, obhajoba metody self-paced reading

Hlavní výhodou této metody je její **dostupnost** – metoda je snadná a levná. První uživatelé této metody museli vytvářet vlastní software pro spuštění experimentu na počítači. Později se však rozšířily standardní „balíčky“ sloužící pro vygenerování experimentů, z nichž např. DMDX je volně ke stažení na internetu. Zároveň je experiment snadno proveditelný na standardním stolním počítači či

notebooku – díky těmto technickým odlišnostem je s to realizovat tyto experimenty mnohem více výzkumníků, jelikož méně laboratoří je vybaveno drahou moderní technikou.

Oproti tomu hlavní kritikou je poukaz (Mitchell, 2004) na **nepřirozenost** metody, na zkreslení „normálních“ operací zahrnutých v běžném zpracování věty. Poukazováno je hlavně na nutnost tisknout klávesu, na nepřirozenou segmentaci předkládaného textu a na zamezení pohledu napřed a zpět v textu během čtení. V rámci obhajoby je namítáno, že tradiční tištěný text byl preferovanou technologií šíření jazyka v rámci jednoho období v historii, postupem času se však v dnešní době rozšiřuje běžné užívání počítačů, kdy mačkání kláves není již ničím nezvyklým, stejně tak jsou široce užívány mobilní telefony, kdy se při čtení SMS zprávy vejde na displej také jen omezené množství několika málo slov.

Dále je zdůrazňováno, že daná metoda zpomaluje proces čtení, který je tudíž kvalitativně odlišný od čtení „přirozeného“, a navíc zpracování textu může být ovlivněno způsobem jakým je předložený text nepřirozeně segmentován. Obě tato tvrzení jsou nepochybně pravdivá. Avšak v prvním případě nebylo ještě nikým dokázáno, že by zveřejněná zjištění byla zcela klamná z důvodu nižší rychlosti čtení. V praxi navíc existuje standardní model, kdy jsou nálezy získané užitím metody self-paced reading potvrzeny užitím dalších (rychlejších) technik jako je např. eyetracking a ERP (Mitchell, 2004, s. 27). Naopak v druhém případě bylo sice dokázáno, že strategie zpracování mohou být ovlivněny konkrétním užitým způsobem segmentace vět. Znovu zde ovšem existuje určitá praxe, kdy pečlivý výzkumník nezveřejňuje závěry před reprodukováním efektu za použití palety odlišných vzorů segmentace – až pokud se závěry potvrdí, lze předpokládat, že to není spojeno s žádným specifickým způsobem členění textu.

Mitchell (2004) tedy tvrdí, že neexistují žádné relevantní důkazy o tom, že by metoda self-paced reading významně komolila proces zpracování vět. Někteří obhájeví této metody si dokonce dovolují tvrdit (Mitchell, 2004, s. 26), že se jedná o pouhou snahu zastánců dražších a komplexnějších metod ospravedlnit jejich odmítání jednoduchých a spolehlivých technik. Navíc tedy, závěry těch jednoduchých byly potvrzeny použitím technik složitých. Obecně lze tedy tvrdit, že co se jeví jedním jako „přirozené“, může se zdát jiným naprosto nepřirozené. Téměř každá technika může nést jak přirozené tak i nepřirozené rysy, důležité tedy

je nedůvěřovat příliš výsledkům, které jsme získali užitím jediné metody, za důvěryhodné nálezy můžeme považovat jen ty, které vzešly z několika experimentů užívajících různé metody.

Existují přirozeně i přijatelná zdůvodnění pro užívání „solistikovanějších“ on-line technik. Ty metody umožňují například analýzu mluveného jazyka (přizpůsobení metod pro sluchovou modalitu) nikoli jen tištěných vět. Navíc poskytují možnost realizace některých metod v situačním kontextu nikoli jen v izolaci ode všeho kromě textu samotného. Poskytují bohatší podobu dat, čímž roste senzitivita úlohy. Zatímco v rámci self-paced reading získáme jen informaci o latencích mezi jednotlivými stisky klávesy, sofistikovnější metody nám poskytují bohatou informaci o povaze problému v daném místě ve větě a ne jen holou informaci o tom, že zde nějaký problém je, tedy že zde latence narostla. A třeba takové neurozobrazovací metody jsou schopné nám dokonce říci nejen to, co se děje, ale i kde jsou dané aktivity lokalizovány v mozku.

I.III.I Vnímání textu při čtení – percepční rozsah (*perceptual span*)

Tím, že užíváme metodu self-paced reading, chceme vyloučit faktor, kdy jedinec při čtení nekouká jen na aktuálně čtené slovo, ale již si prohlíží slova následující, což mu usnadňuje porozumění větám s neobvyklou strukturou, protože se na to jednoduše předem připraví, „předpřečte“ si danou strukturu a ta ho tedy nezaskočí. Metoda self-paced reading mu v tom však brání, čtenář má k dispozici jen jedno slovo, zbytek věty vidí jen ve formě pomlček, což mu však nijak nenapovídá tvar následujícího slova. Z toho vychází jedna ze zmíněných kritik této metody, kdy je napadána přirozenost takového způsobu čtení. Nám jde však o hlubší proniknutí do procesu porozumění větě během čtení, nehodláme naše závěry zobecňovat, ani z nich vyvozovat aplikace pro běžný život, i proto tedy posuzuji metodu jako adekvátní. Přesto však, spíše pro zajímavost, přikládám stručnou zmínku k tematice očních pohybů při čtení (Eysenck, 2008) a velikosti percepčního pole (Rayner, 1998).

Během čtení provádějí naše oči rychlé skoky, tzv. **sakády**. Většinou mají směr dopředu po řádku, občas i zpět, zde hovoříme o tzv. **regresích**. Velikost jedné sakády je asi 8 znaků v textu. Sakády jsou střídány **fixacemi**, kdy se náš pohled zastaví.

Během fixace čerpáme z textu informace potřebné k porozumění čtenému. Velikost efektivního zorného pole v rámci jedné fixace čili **percepční rozsah** je 3-4 písmena (znaky) doleva a 14-15 písmen (znaků) doprava od místa fixace (Rayner, 1998; Eysenck, 2008). Fixace je tedy asymetrická doprava, což je však specifické podle příslušných kulturních podmínek (podle konkrétního systému pravopisu; jinak je tomu u jazyků, kde se text čte zprava doleva). Rozsah je nadále ovlivněn i dalšími faktory jako je vliv dosavadních čtenářských dovedností, podle obtížnosti textu, podle zhuštěnosti textu a otázkou je, zda má vliv to, zda se jedná o slova či jen nesmyslný řetězec písmen. Stejně tak je mezi výzkumníky zájem o to, zda čtenář efektivně sleduje i následující řádek. S tím souvisí i to, že někteří autoři (Rayner a Serenová, 1998, in Eysenck, 2008) rozlišují celkový percepční rozsah, tedy oblast, ze které je člověk schopen extrahovat užitečnou informaci, rozsah identifikace písmen a rozsah identifikace slov, který je podle Raynera (1998) velký jen 7-8 znaků napravo od místa fixace. Dochází zde k tomu, že jsme schopni pracovat s informacemi z parafoveální oblasti na sítnici, což zmenšuje dobu čtení příslušného následujícího slova, nepotřebujeme tolik času na jeho fixaci, neznamena to však, že by zpracování parafoveálních informací dosáhlo sémantické úrovně (Eysenck, 2008).

I.IV Metoda syntaktického primingu

Zmíněný experiment realizovaný na Psychologickém ústavu Akademie věd ČR je zaměřen na sledování přítomnosti jevu tzv. syntaktického primingu – na to, zda syntaktický priming usnadňuje porozumění větám, přičemž snadnost porozumění je měřena skrze rychlost čtení. Mimo mnou pozorované jevy budou tedy nasbíraná data i dále užita v akademické praxi, ač já ve své bakalářské práci uvedu jen jejich malou část.

Smolík (2008) **syntaktický priming** vymezuje jako vliv, který má zpracování gramatické struktury jedné věty na zpracování následně prezentovaných nebo produkovaných vět. Poznává, že podle dostupných nálezů a teoretických návrhů nastává syntaktický priming jen u určitého typu vět. Nastává především ve strukturách, které jsou nějakým způsobem obtížné pro zpracování jako jsou právě víceznačné věty. Vliv syntaktického primingu (Smolík, 2008) je dobře doložený na

produkcí, tedy na vytváření a formulaci vět. Zda se objevuje i v rámci recepce čili porozumění, zjišťuje například i náš experiment, který se zabývá vlivem syntaktického primingu na čtení víceznačných vět, ve kterých na klíčových místech dochází k nutnosti reinterpretace a tím ke zpomalenému čtení.

Syntaktický priming a zpracování (víceznačných) vět

Výzkum syntaktického primingu (Smolík, 2008) přispívá k pochopení podstaty naší **znalosti gramatiky** – tedy jakým způsobem mentálně reprezentujeme znalosti strukturních zákonitostí jazyka. Pro určitou gramatickou strukturu si tvoříme abstraktní reprezentace, to je doloženo právě naší schopností přenášet struktury z jedné věty do druhé, aniž by tyto věty obsahovaly shodná slova (lexikální rovina), právě tento proces je podstatou syntaktického primingu. Skrze výzkum syntaktického primingu se pak snažíme zjistit, jaké jsou ony lingvistické struktury, které lidé reprezentují.

Braniganová a Pickering (Smolík, 2008) přicházejí v souvislosti s výzkumy syntaktického primingu s modelem zpracování gramatické struktury vět vycházející z aktivace tzv. **kombinatorických uzlů**. (Zde se tematicky navracím do oblasti teorií zpracování (víceznačných) vět.) Tyto kombinatorické uzly tvoří reprezentační jednotky zastupující určité kombinace a sledy gramatických kategorií. Uzly jsou aktivovány pokaždé, když jsou tyto kombinace či sledy použity v produkci nebo když se objeví v materiálu určenému k porozumění. Je-li tedy při produkci jedné věty aktivován určitý uzel, pak při produkci další věty je snazší znovu použít tuto konkrétní bezprostředně užitou strukturu. Míra aktivace se snižuje postupně, při produkci dalších a dalších vět je tedy pravděpodobné, že uijeme ten více aktivovaný uzel, s počtem vět se pravděpodobnost užití daného uzlu snižuje s tím jak klesá míra jeho aktivace. Aplikujeme-li tuto teorii i na porozumění, pak čtení věty by tedy mělo být rychlejší, pokud jí předcházela věta se stejnou gramatickou strukturou.

Naším předpokladem v experimentu je, že pokud jedinec přečte jednu větu, kde se setká s danou strukturou, v následující větě už je na tuto strukturu připravenější, než kdyby se s danou strukturou bezprostředně předtím nesetkal. Předložíme jedinci k přečtení větu (tzv. *prime*), která má určitou strukturu, věta je navíc postavena významově tak, že zde jedince užitá struktura nepřekvapí, je běžná,

obvyklá. Poté mu předložíme k přečtení větu, která má stejnou strukturu, ale z hlediska obsahu věty je tato struktura méně obvyklá. Předpokládáme, že v tomto případě bude jedinec méně překvapen než kdyby předcházející věta měla strukturu zcela jinou, odlišnou než věta neobvyklá. Pro porovnání užijeme data získaná v druhém případě, kdy předchozí věta nebude mít shodnou strukturu, porozumění neobvyklé větě tedy nepomůže to, že se s danou strukturou setká osoba v bezprostředně předcházející větě.

Předtím, než přikročím k praktické části práce, bych ještě ráda shrnula, co všechno se čtenář dozvěděl v teoretické části. Na jejím počátku velmi stručně zmiňuji **zpracování vět obecně**, konkrétně vztah syntaktické a sémantické analýzy věty. Dále se pak podrobněji věnuji **víceznačným větám**, seznamuji čtenáře s možnými pohledy na víceznačnost a nejvíce prostoru následně náleží teoriím vysvětlujícím možné způsoby interpretace víceznačných vět v rámci procesu porozumění. Hlavními jsou zde **model slepé koleje**, **teorie omezujících faktorů** a **teorie kapacity**. Druhá polovina teoretické části náleží online metodám se zaměřením na **metodu self-paced reading**. V závěru teoretické části práce pak nalezneme čtenář ještě několik málo informací o jevu zvaném **syntaktický priming**, s nímž souvisí i další možná teorie zpracování víceznačných vět vycházející z aktivace tzv. **kombinatorických uzlů**.

Praktická část – příprava a realizace experimentu

II.I	Vymýšlení vět	25
II.I.I	Věty A	25
II.I.II	Věty B	27
II.II	Sestavení dotazníků A a B	28
II.III	Pilotní výzkum	30
II.IV	Příprava samotného experimentu – počítačový program	33
II.V	Administrace počítačového experimentu	33
II.VI	Statistické zpracování dat	35

V experimentu realizovaném metodou self-paced reading hodnotíme porozumění větě skrze rychlost čtení. Nezajímáme se přitom pouze o celkovou dobu, za kterou čtenář větu přečte, ale v souvislosti se srozumitelností víceznačných vět je pro nás zajímavá i plynulost čtení. Předpokládáme, že čtení jednoznačných vět je víceméně plynulé, oproti tomu rychlost čtení vět víceznačných se v průběhu výrazně mění. Jedinec čte jednotlivá slova, ze kterých se skládá věta. Předpokládáme, že jedinec očekává určitou větnou strukturu, má jistou představu, jak bude daná věta pokračovat, ani ne tak z hlediska obsahu (jaké konkrétní slovo bude následovat), ale z hlediska struktury (např. jaký větný člen či jaký slovní druh bude následovat, jaký bude mít příští slovo pád). Myslíme si, že jedinec toto očekávání zakládá na zkušenosti, případně na kontextu, který z věty vyplývá. Předpokládá něco, co je obvyklé, běžné, my se však snažíme jeho očekávání narušit. Zvolíme strukturu, kterou považujeme za méně obvyklou v běžné řeči. Předpokládáme, že pokud předložíme k přečtení větu, která má neobvyklou strukturu, jedinec se při jejím čtení „zarazí“. U tohoto slova (či tvaru slova) se podle nás tempo čtení zpomalí, což považujeme za ukazatel narušeného porozumění větě. Jedinec je překvapen danou strukturou, přeruší se plynulost čtení, protože se musí „zamyslet“ nad smyslem čtené věty. Jeho očekávání běžně napomáhá porozumění větě, my však jeho očekávání záměrně zklameme.

Z počáteční úvahy o zaměření mé bakalářské práce uvedené v „Úvodu“, vyplývají následující **hypotézy** a (mnou uvažované) alternativy řešení:

(H1) *Je čtení vět víceznačných pomalejší než čtení obdobných vět jednoznačných?*

(H2) *Existuje závislost (korelace) mezi dobou čtení vět víceznačných a hodnocením jejich srozumitelnosti?*

(H3) *Promítá se do této závislosti (viz. H2) také počet znaků ve větě?*

Předpokládejme, že mé **alternativy** tvoří dva krajní body na škále (dvě opačné možnosti), kdy reálné výsledky se mohou nacházet někde na této „přímce“. Pokud jde o otázky/hypotézy k mé bakalářské práci, v prvním případě (H1) jsou alternativy následující: Předpokládám, že (průměrná) doba čtení (předložených) vět víceznačných je delší než (průměrná) doba čtení příslušných vět jednoznačných, je však možné, že to bude i naopak, tedy že delší doba čtení bude na straně vět jednoznačných, případně, že bude doba čtení vět víceznačných i jednoznačných stejná. V druhém případě (H2) předpokládám, že existuje závislost mezi dobou čtení vět víceznačných a hodnocením jejich srozumitelnosti, navíc, že je tato závislost přímá, tedy taková, kdy u vět s horším hodnocením srozumitelnosti je zároveň delší doba čtení. Dále se ještě domnívám (H3), že na dobu čtení má vliv i počet znaků ve větě. V případě potvrzení námi předpokládané závislosti by tedy věta s více znaky byla hodnocena jako hůře srozumitelná a měla by zároveň delší dobu čtení.

Praktická část je z hlediska užitých metod rozdělena do několika kroků. V první řadě bylo potřeba vytvořit soubory víceznačných vět, se kterými bylo dále v experimentu pracováno. Věty byly posléze zpracovány do podoby dotazníků a zadány studentům, kteří měli za úkol ohodnotit jejich srozumitelnost v prvním případě a doplnit preferovaný tvar spojení v druhém případě. Poté nastala hlavní část experimentu, kdy byly věty transformovány do podoby počítačového experimentu a ten byl zadán pokusným osobám, které měly za úkol přečíst předložené věty metodou self-paced reading. Všechny tyto kroky sloužily k nashromáždění potřebných dat. Posledním krokem v této posloupnosti pak tedy bylo provést jejich statistickou analýzu.

II.I Vymýšlení vět

Z nepřehledného množství víceznačných vět jsem měla za úkol vytvořit dvě sady, kdy každá byla tvořena jedním typem víceznačných vět. Pracovalo se s přechodně víceznačnými větami, což znamená, že věty jsou během čtení zpočátku víceznačné, ale jakmile jedinec dočte větu do konce (nebo do určitého místa), smysl věty jednoznačně pochopí. Víceznačnost je navíc daná slovním tvarem nikoli významově, proto hovoříme o gramatické víceznačnosti.

II.I.I Věty A

U prvního typu vět jsme si „pohráli“ s přechodnou víceznačností, která je způsobena tvarem slova v určitém pádu. Využili jsme toho, že tvary některých slov jsou shodné ve více pádech – tedy, čtenáři není v prvním okamžiku zcela jasné v jakém je dané slovo pádu, což ovlivňuje jeho porozumění celé větě. Předpokladem je, že čtenář se zde během čtení věty opírá ve snaze o porozumění obsahu o svou zkušenost s českým jazykem – že si jedinec v průběhu čtení myslí, že je slovo v pádu, který je z hlediska struktury věty běžnější, ale když přečte následující větný člen, zjistí, že se jednalo o pád jiný. Zároveň jsou věty postaveny významově tak, aby čtenáře naváděly k tomuto chybnému počátečnímu vnímání pádu daného slova, až následně zjistí, že se jednalo o pád jiný. Právě u toho následujícího slova, které prozradí s ohledem na rozšířený smysl věty skutečný pád slova předchozího, předpokládáme zpomalení plynulého čtení, protože se tím naruší čtenářovo očekávání. Myslí si, že věta hovoří o něčem a předpokládá jak zhruba bude pokračovat, najednou ale zjistí, že jeho úvahy jsou chybné a věta má ve skutečnosti trochu jiný význam (obsahový) – např. „myslivec zastřelil prasnici sele“ – myslivec nezastřelil prasnici, jak by se zprvu mohlo zdát, ale zastřelil sele, které patřilo prasnici, jak zjišťujeme vzápětí během čtení.

Konkrétně tu užíváme slov (podstatných jmen), u kterých se shoduje tvar ve třetím a čtvrtém pádu. Věty pak tvoříme záměrně tak, aby byl vzbuzen dojem, že se jedná o 4. pád, ale následně čtenář z kontextu (z významu věty) zjišťuje, že se jednalo o pád třetí. Užitá větná struktura vypadá tedy následovně:

Podmět (aktér) + **přísudek** (děj) + **předmět** (3. pád, *Komu/čemu?*)
+ **předmět** (4. pád, *Koho/co?*) + **příslovečné určení**
(v předložkovém tvaru, *Kde?, Kdy?, Proč?, Jak?*).

Věta obsahuje **podmět**, který vyjadřuje činitele děje, „někdo“ *něco dělá*, **přísudek**, který pojmenovává děj, „co“ *ten někdo dělá*, **předmět ve 3. pádu**, jehož tvar může být zároveň 4. pádem, navíc ve spojení s přísudkem čtenáři „podsouváme“, že se o 4. pád jedná, předmět vyjadřuje (*ke*) „komu/čemu“ *patří ten či to, kterému se něco děje*, následuje **předmět ve 4. pádu**, ze kterého se dozvídáme „koho/co“ *se děj týká, tedy komu je to něco děláno* – ze smyslu věty v tomto okamžiku tedy vyplývá, že první předmět byl v pádu třetím, věta je dokončena **příslovečným určením** času, místa, způsobu či příčiny, přičemž užíváme tvar s předložkou, tento člen už jen doplňuje okolnosti děje, funkce, kterou zde zastává, je ta, aby slovo, u kterého nás doba čtení zajímá nejvíce, nestálo na konci věty.

První předmět ve větě je, jak už bylo výše zmíněno, podstatné jméno, které má stejný tvar ve 3. i 4. pádu. Tento požadavek však splňuje jen poměrně úzká skupina podstatných jmen. Při tvorbě experimentálních vět jsem užívala podstatných jmen rodu ženského, která se skloňují podle vzoru růže a která mají koncovku (převážně) -ice, -yně, případně -e, nebo podstatných jmen rodu středního skloňovaných podle vzoru stavení, zde však bylo obtížnější (podobně jako u rodu ženského s koncovkou -e) vymyslet spojení, které by z hlediska významu napovídalo čtenáři na spojení přísudku s předmětem ve zdánlivém 4. pádu. Seznam konkrétních vět použitých v experimentu najdete v Příloze 1.

II.I.II Věty B

V druhém případě se vycházelo z vět, jejichž základem bylo slovní spojení podmětu s přísudkem. Z lingvistického hlediska jsme zde pracovali s valencí slovesa a s vlastností zvanou tranzitivita, kdy přísudek vyžaduje valenci předmětu ve 4. pádu. Pro ilustraci uvedu např. sloveso „pršet“, ke kterému předmět ve 4. pádu připojit nemůžeme. Pak existují slovesa jako třeba „hrát“, které např. ve spojení „Pepa hraje“ vyžaduje nějaké doplnění, aby věta dávala smysl. A nakonec existují slovesa tzv. obojetná, jako třeba „snídat“, kdy můžeme vytvořit jednoduchou větu „Pepa snídá“, která dává smysl aniž bychom museli připojit předmět, ale stejně tak můžeme vytvořit větu „Pepa snídá housku“, kde ke slovesu připojíme předmět ve 4. pádu a věta stále působí přirozeně, nešroubovaně.

V našem případě jsme se zaměřili konkrétně na věty, ve kterých je přísudek tvořen slovesem obojetným – tedy tím, které dává smysl samo o sobě, ale zároveň může být doplněno předmětem ve 4. pádu. V rámci experimentu se pak pracovalo s preferencí tranzitivních či intranzitivních spojení, tedy zda čtenář preferuje u obojetného slovesa doplnění předmětu ve 4. pádu (tranzitivní spojení) nebo něčeho jiného, v našich větách převážně příslovečných určení (intranitivní spojení). Opět jsme vycházeli z určitého předpokladu, že čtenáři preferují spojení tranzitivní. Tudíž nás zajímalo, zda, pokud narušíme předpokládané očekávání a předložíme čtenáři spojení intranzitivní, zpomalí se jeho doba čtení, protože bude předloženou gramatickou strukturou zaskočen. Údaje z pilotního výzkumu pak měly sloužit pro potvrzení naší domněnky, že čtenáři preferují u daných sloves spojení tranzitivní. Seznam konkrétních vět použitých v experimentu najdete v Příloze 2.

Podmět (aktér) + přísudek (děj vyjádřený obojetným slovesem)
+ in/tranzitivní dokončení věty.

II.II Sestavení dotazníků A a B

Jako první byly věty zpracovány do podoby dvou dotazníků, kde jsme v prvním případě (tedy u vět obsahujících záměnnost 3. a 4. pádu) zjišťovali jejich srozumitelnost skrze výběr číselné hodnoty na sedmibodové škále. V druhém případě (preferenze tranzitivního či intransitivního spojení) bylo úkolem respondentů doplnit spojení tvořené podmětem a přísudkem libovolně tak, aby dokončili větu.

Dotazník A

Jak jsem popsala výše, mým úkolem bylo vytvořit sadu vět A, celkem jich bylo 35. K těm přibylo dalších 31 vět, které dodal konzultant z Psychologického ústavu AV ČR. I tyto věty se tedy následně staly součástí počítačového experimentu. Opět se jednalo o věty víceznačné, jejich seznam naleznete v Příloze 3. Stejně jako u „mých“ vět se i u těch ostatních sledovalo, jak jejich specifická gramatická struktura ovlivní čtenářovo porozumění. Celkový soubor tak čítal 66 víceznačných vět.

Pomocí tabulkového editoru byla vygenerována čtyři různá náhodná pořadí všech vět. Následně byly věty přeneseny do textového editoru, kde jsem vytvořila finální podobu dotazníků. Ta je k nahlédnutí v Příloze 4. Dotazníky tedy existují ve 4 verzích, které se liší pouze pořadím vět, které je, jak jsem zmínila, čistě náhodné. To bylo učiněno proto, abychom omezili to, že hodnocení vět respondenty, ovlivní právě pořadí vět.

Kromě víceznačných vět obsahuje dotazník i stručnou **instrukci**, která se opakuje na počátku každé stránky – to proto, abychom minimalizovali šum při otáčení stránek, kdyby si chtěli respondenti instrukci připomenout, takhle ji mají tedy neustále k dispozici a není potřeba listovat na začátek dotazníku. Hlavní náplň dotazníku tvoří jednotlivé **položky**, kdy je každá věta bezprostředně doplněna o sedmibodovou číselnou škálu, na které mají respondenti zaznačit, nakolik se jim zdá věta po přečtení srozumitelná. U každé škály je připomenuto, že číslo (1) znamená „velmi snadno srozumitelná“ a číslo (7) „velmi obtížně srozumitelná“, abychom zabránily zkreslení dat v případě, že by si respondent nezapamatoval správně instrukci a údaje prohodil. Jednotlivé položky, kdy položku tvoří věta

a škála, jsou od sebe přehledně odděleny. Tak jako u většiny dotazníků jsme se respondentů ptali i na některé **základní údaje** o jejich osobě, které by se daly využít při statistickém zpracování. Zajímá nás věk, pohlaví a třída, kterou respondent navštěvuje, aby bylo případně možné zhodnotit ovlivnění odpovědí prostředím, ve kterém byly dotazníky vyplněny.

Dotazník B

Dotazník sestavený z vět B, který je k nahlédnutí v Příloze 5, má z většiny totožnou strukturu jako dotazník sestavený z vět A. Oproti dotazníku A byly v dotazníku B použity pouze věty (slovní spojení) B, nic dalšího se zde tedy nedoplňovalo. Položek bylo tudíž jen 40. Opět jsem v první řadě pomocí tabulkového editoru vygenerovala čtyři různá náhodná pořadí vět, jednotlivé soubory vět jsem následně přenesla do textového editoru, kde jsem vytvořila finální podobu dotazníků. Stejně jako u dotazníku A, i zde tedy vznikly čtyři verze dotazníku, které se lišily pouze pořadím vět. I zde jsme chtěli eliminovat možný vliv pořadí vět (slovních spojení) na jejich doplnění.

Dotazníky obsahují stručnou **instrukci**, která je umístěna na počátku každé strany dotazníku, to proto, aby respondent nemusel listovat na začátek, pokud by si v průběhu vyplňování chtěl instrukci připomenout. Oproti dotazníku A je zde instrukce mírně pozměněna, což samozřejmě vyplývá z toho, že zde respondenti pracují s odlišným typem vět a zároveň odlišným úkolem. Zcela totožná je část, která se vyptává na **základní údaje** o osobě jako jsou věk, pohlaví a třída, kterou respondent navštěvuje – opět pro případnou možnost ověřit si, zda za shodnými doplněními nestojí přítomnost vyplňujících respondentů v jedné místnosti. To, v čem se dotazníky odlišují, je charakter položek. Celou **položku** tvoří pouze samotné slovní spojení. Stejně tak jak jsou tato spojení uvedena v Příloze 2, je u každého spojení ponechán vyznačený prostor, kam má respondent za úkol doplnit výraz či výrazy, tak aby dle svého uvážení vhodně doplnil celou větu. Předpokládali jsme zde, že respondenti budou doplňovat především tranzitivní spojení, tedy předměty ve 4. pádě, s menší četností jsme pak očekávali příslovečná určení a předměty v jiných pádech. Navíc jsme jim nabídli „krátký“ prostor pro doplnění, ač jsme je vybídli, že vyznačeným prostorem se nemusí nechat při doplňování omezit,

podvědomě jsme je chtěli přimět k doplněním jednoslovným. Struktura dotazníku zde nevyžaduje optické oddělení jednotlivých položek pro větší přehlednost. Dotazník B je jako celek oproti dotazníku A kratší a vypadá jednodušeji.

II.III Pilotní výzkum

Se žádostí o spolupráci při sběru dat formou dotazníků jsem se obrátila na všeobecné gymnázium v Roudnici nad Labem. Na této škole jsem sama před lety studovala, předpokládala jsem tedy vstřícný přístup vedení školy. Pan ředitel po krátkém uvedení do zkoumané problematiky souhlasil a poslal mě, abych se na čase, který strávím se studenty, dohodla již s konkrétními pedagogy.

Studenty střední školy jsme zvolili ze dvou důvodů. Za prvé z důvodu čistě praktického, kterým byla úspora času. V rámci třídy je možné sesbírat velké množství dat během jednoho časového úseku. Navíc se nám hodilo i z hlediska věku, což byl druhý důvod našeho výběru. Předpokládali jsme totiž, že účastníky počítačového experimentu budou, tak jak tomu bývá na Psychologickém ústavu AV ČR často, studenti VŠ. Mezi oběma skupinami studentů (těmi ze střední i vysoké školy) nejsou větší věkové rozdíly, což je přirozeně lepší, chceme-li výsledky obou skupin posléze porovnávat. Jediným mým požadavkem na pedagogy tedy bylo, aby vybraní studenti navštěvovali tzv. vyšší gymnázium. Výběr samotných tříd už byl pak věcí náhody, podle ochoty a aktuálního rozvrhu oslovených pedagogů. S potěšením jsem využila možnosti nasbírat potřebná data najednou během jediné návštěvy v zařízení.

Do jednotlivých tříd jsem následně vstupovala vždy v doprovodu pedagoga, který studentům oznámil pouze to, že se zúčastní vyplňování dotazníků. Během krátkého představení jsem uvedla, že jsem studentka psychologie a chtěla bych je požádat o spolupráci na výzkumu realizovaném na Psychologickém ústavu Akademie věd ČR.

Dotazníky jsem jim rozdala tak, aby se nevyskytovaly dvě stejné verze vedle sebe. Zároveň jsem je požádala, ať s vyplňováním ještě na chvíli počkají. Následně byli instruováni, aby vyplnili předložený dotazník přesně podle stručné instrukce v hlavičce dotazníku. Instrukce jim byla poté i přečtena a zároveň byli upozorněni, že ji naleznou na začátku každé stránky. U Dotazníku A jsem je požádala, aby si

postupně přečetli jednotlivé věty a ty bezprostředně po přečtení ohodnotili z hlediska jejich srozumitelnosti na nabídnuté sedmibodové škále. Zdůraznila jsem, že mají zaškrtnout jedno ze sedmi čísel, přičemž je u každé škály zopakováno, že hodnota (1) znamená „velmi snadno srozumitelná“ a hodnota (7) „velmi obtížně srozumitelná“. U Dotazníku B jsem je požádala, aby si přečetli jednotlivá spojení a každé podle smyslu libovolně doplnili tak, aby vznikla ucelená věta. Dodala jsem, že není nutné řídit se při doplňování délkou vytečkované čáry. U obou dotazníků jsem následně ještě dodala, že nad úkolem není potřeba déle přemýšlet, postačí, pokud napíšete první co je napadne v případě Dotazníku B, stejně tak ohodnotíte v případě Dotazníku A větu podle prvotního dojmu. Studenti kladli doplňující dotazy, ty však bylo možné zodpovědět zopakováním instrukce, nebylo potřeba dalších doplňujících informací. Padlo i několik otázek ohledně zaměření výzkumu, ale k tomu jsem jen dodala, že na to jim odpovím až po vyplnění dotazníků, chtěla jsem se vyhnout tomu, že by studenty napadlo, co se od nich v dotaznících očekává, především v Dotazníku B. V závěru jsem ještě dodala, aby nezapomněli vyplnit i doplňující údaje o své osobě. Každý student následně vyplnil jeden dotazník. Studenti byli požádáni, aby během vyplňování nekomunikovali, nepokřikovali a ani nenahliželi ke spolužákům. Tento požadavek byl vesměs tolerován, tudíž nebylo potřeba žádný z dotazníků vyloučit pro duplicitu odpovědí.

Vypracování trvalo u Dotazníku A zhruba 20 minut, u Dotazníku B asi 15 minut. Každý z dotazníků vypracovali studenti dvou tříd, celkem se tedy pilotního výzkumu zúčastnily čtyři třídy. Základní charakteristiky zúčastněných osob, jako jsou pohlaví a věk naleznete v následujících tabulkách, podrobnější data se nacházejí v tabulce v Příloze 9.

Tabulka, Dotazník A					
Účastníci pilotního výzkumu, Dotazník A					
Respondenti	N	Věk			
		průměr	min	max	sm. odch.
Ženy	23	17,13	16	18	0,69
Muži	17	17,65	16	19	0,79
Celkem	40	17,35	16	19	0,77

Tabulka, Dotazník B					
Účastníci pilotního výzkumu, Dotazník B					
Respondenti	N	Věk			
		průměr	min	max	sm. odch.
Ženy	28	18,07	16	20	1,21
Muži	15	18,13	16	19	0,99
Celkem	43	18,09	16	20	1,13

Přehled toho, jak jednotliví respondenti v Dotazníku A hodnotili srozumitelnost vět, včetně dalších statistických údajů, je uveden v Příloze 6. U většiny vět bylo hodnocení srozumitelnosti rozloženo v rámci celé sedmibodové škály (jen připomínám, že hodnota 1 označovala větu „velmi snadno srozumitelnou“, zatímco hodnota 7 označovala větu „velmi obtížně srozumitelnou“). U všech vět byla nejmenší (tedy nejlepší) hodnota srozumitelnosti 1. Nejvyšší (tedy nejhorší) hodnota srozumitelnosti dosahovala u většiny vět hodnoty 7, pouze u jedné věty byla nejvyšší (nejhorší) hodnota srozumitelnosti (pouze) 4. Pro zajímavost uvádím v následující tabulce tři nejlépe (nejníže) a nejhůře (nejvýše) hodnocené věty. S daty získanými z Dotazníku B jsem při statistické analýze neoperovala, tudíž je v práci neuvádím.

Tabulka, Přehled nejlépe a nejhůře srozumitelných víceznačných vět			
		Průměrná hodnota srozumitelnosti	Počet znaků ve větě
Nejlépe hodnocené věty			
1	Selka přichystala slepici krmení do korýtka.	1,43	44
2	Vtipálek ostříhal kadeřnici vlasy během spánku.	1,45	47
3	Pán odeslal hospodyni výplatu na konto.	1,65	39
Nejhůře srozumitelné věty			
1	Obžalovaný podplatil porotkyni kolegu před soudem.	4,35	50
2	Vyšetřovatel zatknul zrádkyni syna za krádež.	3,73	45
3	Policista chytil uličnici papouška na ulici.	3,65	44

V předchozích několika tabulkách uvádím jako jednu z charakteristik účastníků jejich **pohlaví**. Toto rozdělení na muže a ženy je vesměs konvenční. Nijak dál s ním nepracuji a ani nechci vznášet domněnku, že by snad existovaly statisticky významné rozdíly mezi hodnocením srozumitelnosti ze stany mužů a žen. Obdobné rozdělení se vyskytuje v tabulkách týkajících se účastníků počítačového experimentu. Opět se jedná pouze o shrnutí údajů, které jsme od jednotlivých

účastníků získali o jejich vlastní osobě. Nepředpokládám rozdíly ve výkonu založené na pohlaví a ani tuto možnost v práci statisticky neověřuji.

II.IV Příprava samotného experimentu – počítačový program

V rámci realizace hlavní části práce byly mé věty, společně s dalšími větami, které dodal konzultant z Psychologického ústavu, zpracovány do podoby počítačového self-paced reading experimentu. Z možných variant této metody jsme v našem experimentu použili tzv. nekumulativní moving-window verzi. Účastníci měli za úkol číst věty „vlastním tempem“ slovo po slově na počítači, zatímco byla měřena doba čtení jednotlivých slov. Z té je pak možno skrze statistické analýzy usuzovat na přítomnost či nepřítomnost různých sledovaných jevů – zde především tedy na vliv syntaktického primingu na porozumění čtenému textu.

Experiment obsahoval celkem 80 položek, přičemž jednu položku tvořila čtveřice vět a příslušná doplňující otázka. Úkolem pokusné osoby bylo přečíst v tichosti co nejrychleji ale zároveň pozorně každou čtveřici a správně zodpovědět následující otázku. První věta ve čtveřici byla tzv. *prime*, tedy věta, která ovlivní nebo neovlivní následně čtenou (víceznačnou) větu. Druhá věta ve čtveřici byla věta cílová, tzv. *target*, tedy věta ovlivňovaná. Právě u této věty nás zajímalo, zda rychlost jejího čtení bude ovlivněna strukturou předcházející věty, tedy primu, která měla buď shodnou či rozdílnou strukturu. Zbylé dvě věty jsou věty výplňkové, tzv. *fillery*. Ty mají za úkol zmást pokusnou osobu, aby neodhalila záměr výzkumu a neovlivnila tak záměrně či nezáměrně jeho výsledek.

Závislou proměnnou je v experimentu čas čtení víceznačné věty, především čas čtení slova, u kterého čtenář víceznačnost odhaluje, kde případně zjišťuje, že jeho počáteční interpretace byla chybná. Zajímají nás zde rozdíly mezi rychlostí čtení těchto úseků u vět, kterým předcházela věta se shodnou či odlišnou strukturou.

II.V Administrace počítačového experimentu

Pokusná osoba seděla u stolu, na kterém byl položen notebook. Nacházela se v klidné a tiché místnosti, kde ji nic nerušilo. Ze stejného důvodu byla požádána,

aby si vypnula mobilní telefon. Před začátkem administrace jsem uvedla osobu velice stručně do problematiky, kdy jsem jí sdělila, že se zúčastní experimentu, který se zabývá porozumění větám, že tento experiment je součástí mé bakalářské práce a úkol, který pro ni mám připravený se skládá z čtení vět a odpovídání na otázky, navíc to celé nezabere více než 30 minut.

Během experimentu viděla osoba na monitoru čtyři pod sebou napsané věty, kdy písmenka byla nahrazena pomlčkami, např. věta *Tatínek obědvá u stolu.* by v experimentu vypadala takto: ----- . Jednotlivá slova osoba postupně odkrývala pomocí stisků mezerníku, vždy však viděla pouze jedno slovo. Po prvním stisknutí se tedy zobrazilo první slovo, při druhém stisknutí se první slovo proměnilo zpět v pomlčky a zobrazilo se druhé slovo a tak dále až k poslednímu slovu. Po co nejrychlejším přečtení 4 předložených vět se po dalším stisknutí mezerníku zobrazila kontrolní otázka, která se vztahovala k jedné ze 4 právě přečtených vět. U příkladové věty by otázka mohla znít např. takto: *Obědvá tatínek v posteli?* nebo *Obědvá tatínek s maminkou?* Osoby byly upozorněny, že pokud se ptá otázka na něco, co se z věty nedozvěděly, odpověď je automaticky záporná. Kontrolní otázka byla mezi věty vložena proto, abychom přiměli čtenáře číst pozorně a vyhnuli se tak situaci, kdy jen prokliká jednotlivá slova, aniž by věděl, o čem věta mluví. Na otázku bylo možné odpovědět jen ANO/NE. Odpovídalo se pomocí stisku pravé či levé klávesy CTRL. Jako nápověda bylo na monitoru pod otázkou nalevo napsáno ANO a napravo NE, což odpovídalo značení kláves, tedy levý control byl označen nápisem ANO, pravý control byl označen nápisem NE. Nápisem byly označeny všechny používané klávesy, tedy i mezerník nesl nápis DÁLE.

Na odpověď byla časová dotace 3 vteřiny. Nutili jsme tak osobu, aby byla na práci koncentrovaná. Po odpovědi (stisknutí klávesy control) následovala zpětná vazba, kdy se na monitoru zobrazil na několik vteřin nápis správně či špatně, osoba se tak bezprostředně dozvěděla, zda byla při odpovídání úspěšná. Případná chyba či vypršení časového limitu mělo osobu motivovat, aby se u následujících položek více soustředila. I to bylo zdůrazněno v instrukci. Zpětná vazba po několika vteřinách sama přešla do nápisu: *Další položku spustíte stisknutím mezerníku.* Prvních 5 položek (tedy 20 vět a 4 otázky) bylo zácvičných, zde se ještě mohla pokusná osoba ptát, pokud jí bylo něco nejasné. Následně nápis mezi položkami upozornil, že další položka je již součástí experimentu, stejně tak se po poslední

položce objevila informace o konci experimentu.

Sdělování instrukce k úkolu jsem pro lepší srozumitelnost doplnila názornou ukázkou, kdy jsem za osobu vypracovala první zácvičnou položku. Znění instrukce je k dispozici v Příloze 8. V rámci důležitých upozornění bych zdůraznila například to, že jsem požádala osobu, aby se pokusila experiment nepřerušit, pokud to nebude nezbytně nutné. Pro chvíle odpočinku mohla využít okamžiku, kdy je mezi položkami na monitoru nápis oznamující *spuštění další položky stisknutím mezerníku*, pouze v tento okamžik se neměřil čas. Upozorňovala jsem, že některé věty mohou znít podivně, neobvykle či dětinsky, a žádala, aby si toho zkusily jedinci nevšímat a nerušeně pokračovat ve čtení. Doporučila jsem připravit si ruce na klávesnici, aby nevznikali zbytečné prodlevy v čase během hledání správné klávesy a stejně tak jsem předvedla jak si představujeme plynulé mačkání mezerníku bez zbytečných prodlev. Všechny zúčastněné osoby již měly předchozí zkušenosti s prací na počítači, což se zřejmě odrazilo na tom, že s vypracováním experimentu nevznikly u nikoho výrazné obtíže.

Pokud jde o výběr osob, jediným požadavkem bylo, aby byla osoba rodilým mluvčím, dále jsem ještě zdůrazňovala, že pokud někdo používá ke čtení brýle, necht' si je nezapomene přinést s sebou i na experiment. Základní charakteristiky (pohlaví a věk) zúčastněných osob, u kterých jsem administrovala experiment já osobně, naleznete v následující tabulce, data od dalších 21 osob dodal konzultant z Psychologického ústavu AV ČR, podrobnější data se nacházejí v tabulce v Příloze 9.

Tabulka, Počítačový self-paced reading experiment					
Účastníci počítačového experimentu					
Pokusné osoby	N	věk			
		průměr	min	max	sm.odch.
Ženy	26	27,15	16	50	8,99
Muži	12	23,75	17	49	8,49
celkem	38	26,08	16	50	8,86

II.VI Statistické zpracování dat a interpretace výsledků

V této části už jsem pracovala jen s vybraným souborem dat. Data byla vybrána s ohledem na zodpovězení mnou na počátku položených hypotéz, ve kterých jsem se zaměřila pouze na věty A (věty se záměnností 3. a 4. pádu). V počítačovém experimentu bylo ze všech vět A použito 32 z nich, jejich seznam doplněný o jejich jednoznačné alternativy naleznete v Příloze 7 a dozvíte se tam i to, která věta tvoří který stimulus („číslo“ věty, pod kterým se následně objevuje ve výsledcích). V rámci experimentu pak byla nasbírána data od 59 osob. Osoba číslo 41 byla z výsledků vyřazena, toto opatření provedl konzultant z Psychologického ústavu AV ČR, který data z počítačového experimentu zpracovával a mě následně poskytl pouze příslušné výstupy v podobě tabulek s nasbíranými daty, která se týkala pouze vět A, ač pokusné osoby se v experimentu setkaly i s dalšími větami.

Každá osoba během vypracování experimentu přečetla 32 položek sestavených z víceznačných vět A a jejich jednoznačných alternativ. Konkrétně 16 vět bylo víceznačných a 16 vět bylo jednoznačných (jednalo se o příslušné dvojice ke zbývajícím 16ti víceznačným větám z celkového počtu 32 dvojic). První polovina osob (v seznamu výsledků) tedy četla první polovinu víceznačných vět (prvních 16 vět, tedy věty 1 až 16) a druhou polovinu vět jednoznačných (druhých 16 vět, tedy věty 17 až 32). Druhá polovina osob četla druhou polovinu vět víceznačných (druhých 16 vět) a první polovinu vět jednoznačných (prvních 16 vět). Přičemž první polovina víceznačných vět (prvních 16 víceznačných vět) tvořila dvojice s první polovinou vět jednoznačných (s prvními 16ti jednoznačnými větami), druhá polovina byla uskupena opačně. Pořadí vět, ve kterém je pokusná osoba četla, záleželo na verzi experimentu, která na ni vyšla. Střídaly se celkem 4 verze experimentu s konkrétní skladbou vět, avšak i přesto bylo pro každou osobu navíc generováno náhodné pořadí jednotlivých položek.

Hypotéza č. 1

(H1) Je čtení vět víceznačných pomalejší než čtení obdobných vět jednoznačných?

Za účelem zodpovězení této otázky jsem pracovala se dvěma následujícími soubory dat, z nichž by i každý zvlášť sloužil k zodpovězení položené otázky. O to zajímavější je, zda dosáhneme stejného závěru u obou souborů. Jen bych zdůraznila, že v obou případech pracuji pouze s **větami A**, resp. s *průměrnou dobou čtení vět A získanou z počítačového experimentu*. V jednom případě se koukám na dobu čtení (jednoznačných i víceznačných) vět A z hlediska jednotlivých čtenářů, v druhém případě se dívám na dobu čtení (jednoznačných i víceznačných) vět A z hlediska jednotlivých stimulů.

Výpočet z hlediska čtenářů, Varianta „čtenáři“

Průměrná doba čtení všech vět víceznačných (V0) a průměrná doba čtení všech vět jednoznačných (V1) u jednotlivých osob.

Zde jsem pracovala s dvěma soubory dat, kdy jsem si vzala seznam jednotlivých subjektů (59 osob) a zaměřila jsem pozornost na průměrné doby, za které jednotlivé osoby přečetly všechny (své) víceznačné (V0) a jednoznačné (V1) věty. Konkrétní data použitá při výpočtu jsou k nahlédnutí v Příloze 10.

Výpočet z hlediska stimulů, Varianta „stimuly“

Průměrná doba čtení jednotlivých vět víceznačných (V0) a průměrná doba čtení příslušných vět jednoznačných (V1) (u jednotlivých stimulů).

Zde jsem pracovala s dvěma soubory dat, kdy jsem si vzala seznam jednotlivých stimulů (32 dvojic vět) a zaměřila jsem pozornost na průměrné doby, za které všechny osoby přečetly u konkrétního stimulu příslušnou větu víceznačnou (V0) a příslušnou větu jednoznačnou (V1). Konkrétní data použitá při výpočtu jsou k nahlédnutí v Příloze 11.

Testování hypotézy (H1)

Pro otestování vyslovené hypotézy (H1) jsem zvolila postup skládajícího se z několika posloupných kroků (Škaloudová, 1998).

Krok 1: Vzhledem k tomu, že data pro věty V0 a věty V1 je možno v obou variantách spárovat (každá dvojice se týká jednoho subjektu resp. jednoho stimulu), volím v obou případech **t-test pro dva závislé soubory**.

Krok 2: Nulová hypotéza **H0** je shodná pro obě varianty a zní: *Není rozdíl mezi průměrnou dobou čtení vět víceznačných (V0) a vět jednoznačných (V1)*.
Odpovídající alternativní hypotéza **H1** u obou variant zní: *Průměrná doba čtení vět jednoznačných (V1) není stejná jako průměrná doba čtení vět víceznačných (V0)*.

Krok 3: S ohledem na data, která mám k dispozici, a tomu odpovídající znění alternativní hypotézy se v obou případech jedná o **dvoustranný test**.

Krok 4: Výpočet **pozorované hladiny významnosti:**

Varianta „čtenáři“ ... $PHV = 0,00098$

Varianta „stimuly“ ... $PHV = 0,00172$

Krok 5: Vztah pozorované hladiny významnosti a **hladiny významnosti α :**

Varianta „čtenáři“ ... $0,00098 < 0,01$

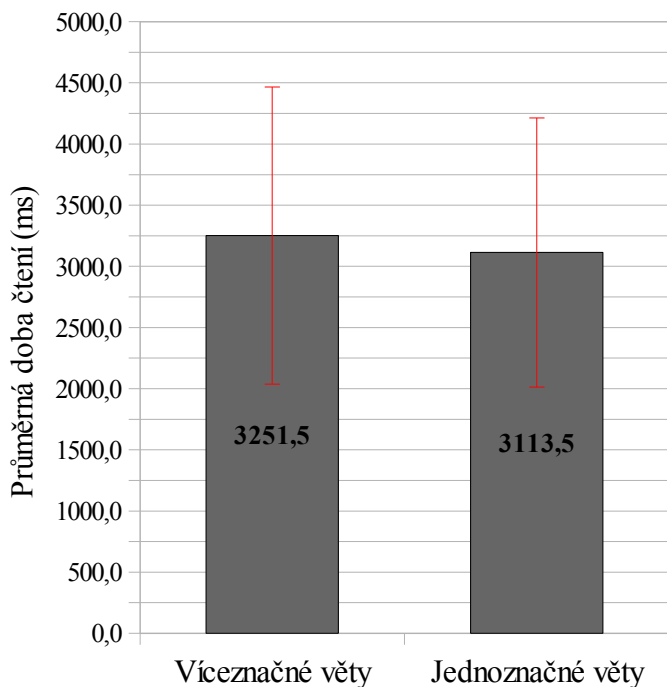
Varianta „stimuly“ ... $0,00172 < 0,01$

Krok 6: Na základě předchozího porovnání v obou variantách H0 zamítám.

Na hladině významnosti 1% bylo (v obou případech) prokázáno, že existuje statisticky významný rozdíl mezi průměrnou dobou čtení vět víceznačných (V0) a vět jednoznačných (V1). Výsledky jsou znázorněny v následujících grafech (viz. Graf zobrazující celkovou průměrnou dobu čtení vět víceznačných (V0) a jednoznačných (V1), Varianta „čtenáři“ a Varianta „stimuly“). Na základě grafického znázornění se mohu vyjádřit i o tom, že doba čtení vět jednoznačných (V1) je u obou variant kratší než doba čtení vět víceznačných (V0).

Varianta "čtenáři"

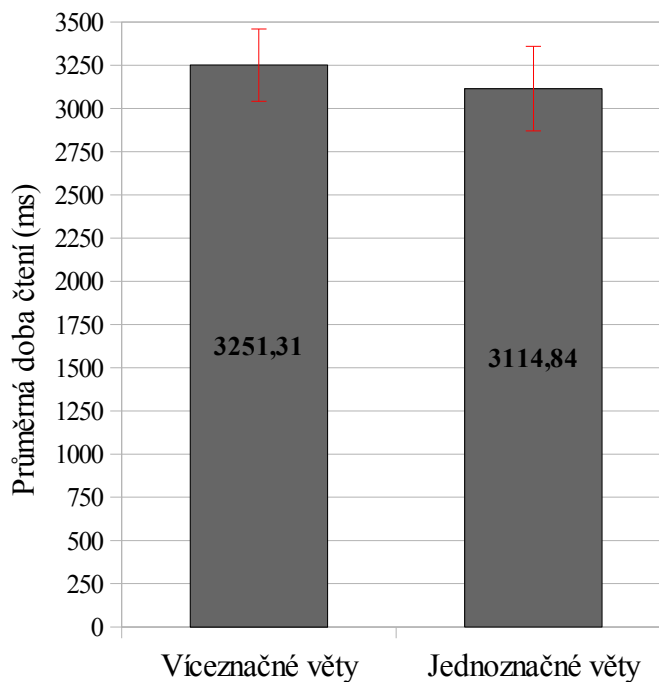
Graf zobrazující celkovou průměrnou dobu čtení vět víceznačných (V0) a jednoznačných (V1)



I ... směrodatná odchylka

Varianta "stimuly"

Graf zobrazující celkovou průměrnou dobu čtení vět víceznačných (V0) a jednoznačných (V1)



Hypotéza č. 2

(H2) Existuje závislost (korelace) mezi dobou čtení vět víceznačných (V0) a hodnocením jejich srozumitelnosti?

Korelační analýza

Pro výpočet korelace jsem použity průměrné doby čtení pro jednotlivé víceznačné věty (V0) a průměrné hodnoty z hodnocení srozumitelnosti jednotlivých V0. Konkrétní data použita při výpočtu jsou k nahlédnutí v Příloze 12.

Korelační koeficient $r = 0,1266$; $N = 32$; $p = 0,4898$

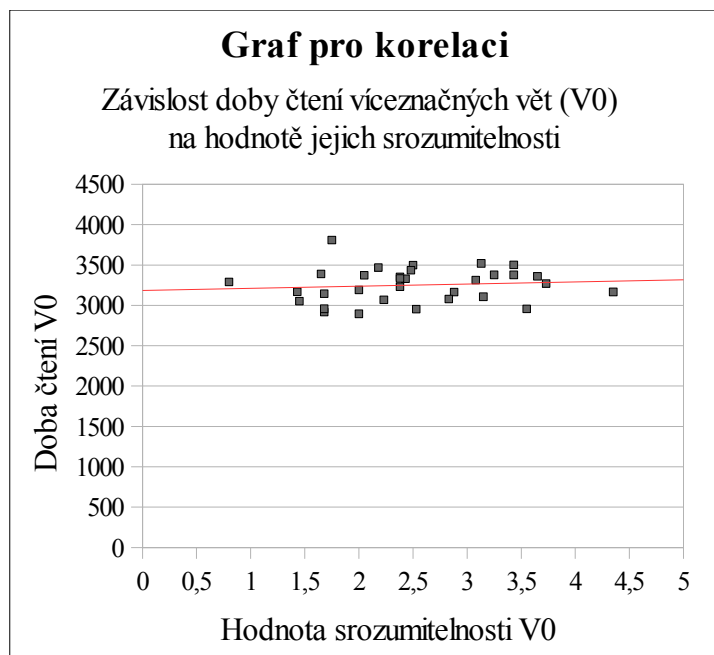
H0: Korelační koeficient je roven 0, tedy neexistuje závislost.

H1: Korelační koeficient není roven 0, tedy existuje závislost.

$\alpha (0,01 \text{ resp. } 0,05) < PHV$

H0 nezamítám: Korelační koeficient je roven 0.

Na základě výpočtu jsem dostala hodnotu korelačního koeficientu $r = 0,1266$ se statistickou významností $p = 0,4898$. S ohledem na blízkost korelačního koeficientu 0 lze na hladině významnosti 1 % říci, že doba čtení a hodnota srozumitelnosti nejsou v případě námi sledovaných vět korelované. Tento závěr je patrný i z následujícího grafu (viz. Graf pro korelaci, Závislost doby čtení víceznačných vět (V0) na hodnotě jejich srozumitelnosti).



— ... regresní přímka

Hypotéza č. 3

(H3) Promítá se do této závislosti (viz. H2) také počet znaků ve větě?

Do statistické analýzy doby čtení víceznačných vět jsem zařadila hypotézu č. 3 (H3) na základě předpokladu, že i délka věty může ovlivňovat dobu jejího čtení. Zároveň však délka věty nemusí odpovídat hodnotě její srozumitelnosti. Je pravděpodobné, že určitá věta obsahující dlouhá slova bude hodnocena jako snadno srozumitelná, a naopak, že nějaká krátká a stručná věta bude z hlediska své srozumitelnosti hodnocena jako obtížná. V následující statistické analýze jsem věnovala pozornost nejprve vztahu mezi dobou čtení věty a počtem jejích znaků izolovaně (korelační analýza), následně jsem se zajímala o existenci kombinovaného vlivu faktorů jako počet znaků a hodnota srozumitelnosti na dobu čtení daných víceznačných vět (vícenásobná lineární regrese). Data použitá pro výpočty jsou k nahlédnutí v Příloze 13.

Korelace

Existuje závislost mezi průměrnou dobou čtení jednotlivých víceznačných vět (V0) a počtem jejich znaků?

Korelační koeficient $r = 0,45$; $N = 32$; $p = 0,0049$

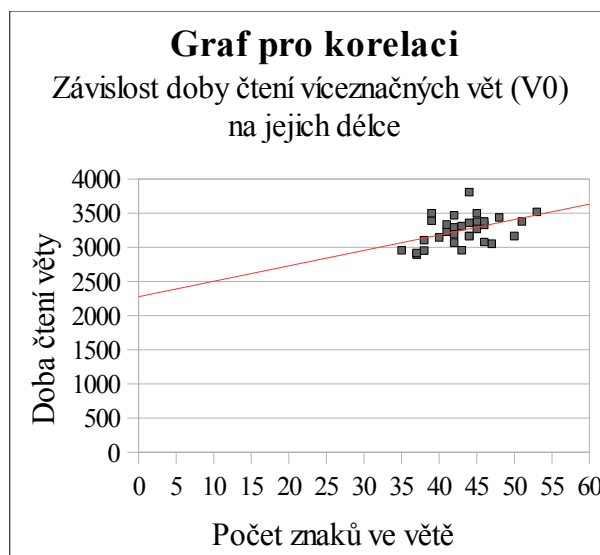
H0: Korelační koeficient je roven 0, tedy závislost neexistuje.

H1: Korelační koeficient není roven 0, závislost existuje.

$\alpha (0,01) > PHV$

H0 zamítám. Korelační koeficient není roven 0.

Na základě základě výpočtu jsem dostala hodnotu korelačního koeficientu $r = 0,45$ se statistickou významností $p = 0,0049$. Na základě těchto hodnot můžu říci, že mezi dobou čtení a počtem znaků ve větě existuje u daných vět mírná pozitivní závislost (Dancey a Reidy, 2002). Proměnné jsou tedy na hladině významnosti 1% korelované, což vidíme i na následujícím grafu (viz. Graf pro korelaci, Závislost doby čtení víceznačných vět (V0) na jejich délce).



— ... regresní přímka

Vícenásobná lineární regrese

Pomocí lineární regrese se snažíme najít funkci, která by nám umožnila predikovat hodnoty závislé proměnné (y) z hodnot jedné či více nezávislých proměnných (x u jednoduché regrese resp. $x_1, x_2 \dots x_n$ u regrese vícenásobné). Hledané hodnoty a a b , které doplníme do rovnice $y = a + b_1x_1 + b_2x_2$, vyčteme z počítačového výstupu některého z programů pro statistické výpočty (např. SPSS), kde dohledáme i příslušné hodnoty statistické významnosti, která nám poví, nakolik přesně sedí nalezená rovnice pro predikci námi sebraných dat.

V našem případě:

y = závislá proměnná, konkrétně doba čtení V0

prediktory = nezávislé proměnné: x_1, x_2

x_1 = počet znaků ve větě

x_2 = hodnota srozumitelnosti V0

$b_1, b_2 \dots$ koeficienty

$a \dots$ konstanta

Vybraná data z počítačového výstupu:

korelační koeficient (dohromady pro x_1 a x_2) $R = 0,454$

statistická významnost **Sig.** = 0,035

	Koeficienty		t	Sig.
	B	Sm.odch.		
Konstanta (a)	2265,679	361,554	6,266	0,000
Počet znaků (x_1)	23,956	9,097	2,633	0,013
Hodnota srozumitelnosti (x_2)	-19,077	50,497	-0,378	0,708

Obecná rovnice vícenásobné lineární regrese: $y = a + b_1x_1 + b_2x_2$

Regresní rovnice: $y = 2265,679 + 23,956 * x_1 - 19,077 * x_2$

H0: Hodnotu y nelze predikovat pomocí příslušné regresní rovnice.

H1: Hodnotu y lze predikovat pomocí příslušné regresní rovnice.

$\alpha (0,05) > \text{Sig.} (0,035) \dots$ tedy: Na hladině významnosti 5 % H0 zamítám.

Dále dílčí H0: Hodnotu y nelze predikovat pomocí dané nezávislé proměnné.

pro x_1 : $\alpha (0,05) > \text{Sig. } (0,013) \dots H_0 \text{ zamítám}$

pro x_2 : $\alpha (0,05) < \text{Sig. } (0,708) \dots H_0 \text{ nezamítám}$

Podíváme-li se na jednotlivé statistické významnosti příslušející nezávislým proměnným neboli prediktorům doby čtení, které jsme získali při výpočtu lineární regrese, dojdeme ke stejným závěrům jako v předchozích dvou korelačních analýzách. Z toho zjišťujeme, že počet znaků ve větě je lepším prediktorem doby čtení dané věty oproti hodnotě srozumitelnosti, která se ukázala jako s dobou čtení nezávislá. V případě regresní analýzy nás však zajímá vliv obou nezávislých proměnných zároveň, v tom případě nám tedy vypočtená regresní rovnice vychází jako statisticky nepřesná pro predikci doby čtení z nasbíraných hodnot (počet znaků ve větě a hodnota srozumitelnosti věty).

Diskuze

Teorie

V teoretické části práce jsem představila některá možná vysvětlení k tematice zpracování (čtených) víceznačných vět, která jsou k dispozici v odborné literatuře. Nyní bych se na danou problematiku ráda podívala z pozice běžného uživatele jazyka. Zamyslím-li se nad situací, kdy se setkám s víceznačnou větou, ve většině případů mě z počátku vůbec nenapadne, že by se mohlo jednat o víceznačnou větu. Až když mnou očekávaný význam není správný, teprve si uvědomím, že je zde další možnost interpretace. To mě alespoň na vědomé úrovni zpracování víceznačných vět přivádí k **modelu slepé koleje**.

Stejně tak jsem s odkazem na běžnou zkušenost zastáncem **vlivu kontextu**. Myslím, že u většiny vět si člověk ani neuvědomí, že se jednalo o větu víceznačnou. Uvažuje automaticky jen jeden význam, který je patrný na základě kontextu výpovědi, případně volí podle obvyklosti struktury. Chci-li vyjádřit jiný význam, zvolím automaticky jinou strukturu, opět tu obvyklejší, tak aby byl zamýšlený význam jasně patrný. V běžné praxi se tedy dle mého názoru málokdy setkáme s víceznačnými větami, kde by byl správným ten méně obvyklý význam. Skrytý význam podstrkujeme čtenáři spíše v případě slovních hříček, případně v útvarech jako je ironie či sarkazmus. Ostatně i v rámci experimentu, kde byl správný ten méně obvyklý význam dané struktury, pokusné osoby se mnohdy domnívaly, že je toto neobvyklé podání chybné, protože se s ním v praxi běžně neseťkávají, nemají tento vzorec zautomatizovaný pro běžné používání. Při výběru uvažované interpretace by tedy na základě této mé úvahy hrála roli dosavadní **míra aktivace daného strukturního vzorce**.

Hypotézy, užitá data

Na základě teoretických výkladů zpracování víceznačných vět jsem si položila tři hypotézy. U **hypotézy 1** jsem vyšla z teorie, že zpracování víceznačných vět je obecně náročnější na pracovní paměť, kdy čtenář musí uvažovat více možných

interpretací ať už postupně či najednou, na základě toho lze tedy předpokládat, že doba zpracování takových vět bude delší než doba zpracování vět jednoznačných, což se mi následně potvrdilo při statistické analýze, kde jsem porovnávala průměrné doby čtení vět víceznačných a jim odpovídajících vět jednoznačných. U dalších dvou hypotéz jsem po statistickém zpracování dat došla k závěru, který se rozcházel s mými očekáváními. U **hypotézy 2** se mi nepotvrdil předpoklad, že mezi dobou čtení víceznačných vět a příslušnou hodnotou srozumitelnosti věty bude existovat závislost. U **hypotézy 3** jsem nadále očekávala, že případná předchozí závislost bude existovat i v případě, že přidám další proměnnou, kterou bude počet znaků ve větě (tedy délka věty). Možných vysvětlení, proč má očekávání nebyla potvrzena, lze nalézt hned několik, zvažují-li chybu na straně užitých metod a skrze ně získaných dat. Při všech výpočtech jsem použila průměrné hodnoty – jednak u doby čtení pro jednotlivé osoby či pro jednotlivé stimuly (jednoznačné a víceznačné věty), dále průměrné hodnoty srozumitelnosti jednotlivých vět. Z dat nebyly vyloučeny údaje, které by se ze souboru výrazně vymykaly (extrémní maxima a minima), užila jsem všechny nasbírané hodnoty. Oba tyto aspekty mohly konečné výsledky u zmíněných hypotéz ovlivnit. Problematické užití průměrných hodnot jsem si však uvědomila až po vypočtení korelační analýzy u hypotézy 2, kdy se mi nepotvrdil můj předpoklad, což mohlo být samozřejmě ovlivněno i tím, jak respondenti hodnotily srozumitelnost víceznačných vět v Dotazníku A (viz dále). Přesto jsem použila průměrné hodnoty i při výpočtech u hypotézy 3. To, že výsledek nevyšel dle mých představ, samozřejmě neznamená, že je chybný a měl by být tedy z práce vypuštěn. Rozhodnutí použít průměrné hodnoty i v následujících výpočtech vyplynulo ze skutečnosti, že je vhodné použít „obdobná“ data v obou případech (u hypotézy 2 i 3), chci-li výsledky navzájem připodobnit či porovnat. Bylo tedy rozumné použít stejný typ dat, tedy průměry, ve všech výpočtech. Myšlenku zopakovat výpočty s užitím rozšířených dat (nikoli tedy průměrů, ale jednotlivých hodnot) jsem zamítla s ohledem na omezený rozsah bakalářské práce. Nevylučuji však možnost, že i při použití jednotlivých hodnot bych došla ke stejnému závěru jako v případě průměrů.

Jak jsem již zmínila, u hypotézy 2 jsem na základě korelační analýzy zjistila, že doba čtení víceznačných vět nekoresponduje s hodnocením jejich srozumitelnosti, ač podíváme-li se na průměrné hodnoty srozumitelnosti jednotlivých vět (průměrné hodnoty proto, že tyto byly použity i při výpočtu korelace), dostáváme poměrně

pestrou škálu, kdy je mezi nejnižší a nejvyšší hodnotou relativně znatelný rozdíl. Považovali-li respondenti jednotlivé věty v rámci dotazníkového šetření za odlišně srozumitelné, jak to, že se tato odlišná srozumitelnost (odlišná obtížnost porozumění větě) neprojevila i na jejich době čtení? Jednou z odpovědí, která se nabízí, je, že by se možné rozdíly vyrovnaly rozličnou délkou vět, tuto možnost jsem však vyloučila následně v rámci regresní analýzy, kdy byl rozličný počet znaků ve větách odečten, ale ani poté se neobjevila statisticky významná závislost mezi dobou čtení a hodnotou srozumitelnosti, ač závislost mezi dobou čtení a počtem znaků se nám potvrdila. Vynechávám zde možné vysvětlení, o kterém jsem se zmínila již dříve, tedy užití (možná) nevhodných dat v podobě průměrných hodnot, ve kterých se nám ztratí mnohá variabilita, navíc jsem uznala, že jsem ze souboru nevyločila ani výrazné extrémy, které mohly finální průměrné hodnoty značně zkreslit. Podívejme se nyní, zda je možné, že problém je v samotném hodnocení srozumitelnosti vět. V rámci experimentu, kdy jsme měřily dobu čtení vět, jsme předpokládali, že se čtenáři „zarazí“ nad užitou strukturní víceznačností a díky tomu se prodlouží doba čtení, která odrazí neobvyklost předložené interpretace. V rámci dotazníkového šetření je však pravděpodobné, že se respondenti řídili převážně podle významu věty aniž by přihlíželi k víceznačnosti struktury. Sledovali tedy použitá slova a to, nakolik se s jejich kombinací setkávají při běžném užívání jazyka. To by se teoreticky dalo ošetřit přesnějším zadáním instrukce, kdy bychom zdůraznili, aby se respondenti soustředili při hodnocení na neobvyklost struktury spíše než na smysl. Dalšími možnými proměnnými, které mohly ovlivnit výsledky dotazníkového šetření je například skutečnost, že k vyplnění dotazníků byli studenti vyzváni během vyučování, jejich motivací tedy byla spíše skutečnost, že přijdou o část výuky, nikoli zájem zapojit se do experimentu, je tedy možné, že podle toho také k vyplnění dotazníku přistupovali a nevyplnili ho „pochtivě“ podle instrukce. Stejně tak se zde ale objevuje i faktor času, kdy měli studenti neomezenou časovou dotaci na to, aby si větu v klidu přečetli a zamysleli se nad ní, oproti tomu osoby při experimentu byli pod určitým tlakem. Případně mohli v rámci dotazníku porovnávat jednotlivé věty mezi sebou a vybírat následně „ty méně a více podivné“. Tyto faktory se mohou odrazit ve faktu, že nám nekoresponduje obtížnost porozumění větám, kterou hodnotíme jednou pomocí škály v dotazníku a jednou na základě doby čtení vět v počítačovém experimentu. Dalším vysvětlením proč jsme nenašli závislost mezi

dobou čtení a hodnocením srozumitelnosti věty je možnost, že věty jsou na obdobné úrovni obtížnosti, možná by bylo vhodnější vymyslet takové věty, aby některé byly opravdu hodně obtížně srozumitelné a některé naopak velmi snadno srozumitelné. Jsou-li si věty v úrovni srozumitelnosti podobné, je tak možné, že závislost nenajdeme, protože doba čtení je v tomto případě ovlivněna jinými faktory.

Metoda

Celková doba čtení věty vs. analýza doby čtení jednotlivých jejích úseků

Již v rámci teoretické části práce zmiňuji jako jednu z kritik **metody self-paced reading** její nepřirozenost, která poukazuje na skutečnost, že čtenář při čtení věty danou metodou vidí vždy jen jedno slovo. Aby dočetl až na konec věty, musí absolvovat několik stisků klávesy. Ústřední myšlenkou metody je možnost měřit čas čtení jednotlivých slov, která jsou „oddělena“ právě jednotlivými stisky. Zde je však riziko, že klikání se stane po chvíli automatickým, riziko roste s počtem přečtených položek. Člověk se pak následně rytmicky „prokliká“ větou, aniž by jednotlivé časové úseky odkazovaly na obtížnost čtení věty v daném místě. Ovšem předpokládáme-li, že člověk i přes automatické klikání zobrazená slova čte, jelikož si je vědom, že bude následně dotázán na obsah věty, je pravděpodobné, že se nad jejím „podivným“ zněním zamyslí dodatečně na konci celé („proklikané“) věty. Potom by celková délka čtení věty včetně doby přemýšlení nad jejím významem mohla nakonec vykazovat obtížnost jejímu porozumění. Zde bychom tedy mohly najít obhajobu pro užití celkového času čtení věty při výpočtech oproti analýze doby čtení jednotlivých úseků, ač právě na tom staví metoda self-paced reading.

Jednou z možností jak přesněji sledovat počínání pokusné osoby napříč větou je užití eyetrackingu, vidíme kam se vrací, jakému slovu či úseku věty při promýšlení významu věnuje zvýšenou pozornost. Výkon sám o sobě je poměrně náročný na pozornost, je tedy nezbytné, aby byl experiment administrován v tiché a klidné místnosti, kde nebude pokusná osoba ničím rušena, zároveň by měla být v „dobré kondici“, čímž myslím, že by neměla být oslabena nemocí či pod vlivem látek, které ovlivňují schopnost soustředění. Předpokládejme, že náš výběr pokusných osob byl důkladný a osoby přistupují k úkolu zodpovědně a vykonávají ho podle naší

instrukce, i přesto je tu vždy riziko, že osoby v rámci věty jen klikají a obsah nečtou, či ho nečtou dostatečně pečlivě, tak aby získané údaje odpovídali obtížnosti porozumění (resp. pozastavení se nad „podivným“ či neočekávaným vývojem obsahu věty). Toto můžeme ošetřit právě důkladnou kontrolou znalosti obsahu textu skrze kontrolní otázky a následnou analýzou chybovosti při odpovídání.

Uvedla jsem zde několik možností, jak „vylepšit“ použití metod či samotné statistické výpočty užitím vhodnějších dat. V souvislosti s tím jsem zmínila některá „slabá místa“ ve své práci. Všechny poukazy jsem ovšem vyslovila pouze v teoretické rovině a nemám pro ně důkazy v podobě praktického ověření. Metody a získaná data, tak jak jsem s nimi pracovala v rámci své práce, jsem použila tak, jak mi bylo doporučeno vedoucím práce a konzultantem z Psychologického ústavu AV ČR. I přes možné chyby mohu s čistým svědomím prohlásit, že zde nedošlo k žádné záměrné manipulaci, která by se snažila změnit výsledky tak aby odpovídaly teoretickým předpokladům.

Závěrečné shrnutí

Pro svou bakalářskou práci jsem si zvolila téma **Porozumění přechodně víceznačným větám**. V praktické rovině jsem se věnovala **přípravě a realizaci psychologického experimentu**.

Využila jsem příležitosti zapojit se do experimentu, který se realizoval na oddělení kognitivní psychologie **Psychologického ústavu AV ČR v Praze**. V rámci tohoto experimentu šlo o manipulaci s experimentálním materiálem, který se skládal z několika typů víceznačných vět, kdy bylo hlavním zájmem experimentátorů zjistit, zda lze u některého z daných typů víceznačných vět vysledovat tzv. **jev syntaktického primingu**. Má spolupráce spočívala ve vymýšlení konkrétních experimentálních vět, které reprezentovali dva typy gramatické víceznačnosti ve větách, a to přesně dle zadané šablony pro určitou větnou skladbu. V dalším kroku jsem věty transformované do podoby dvou dotazníků zadala studentům víceletého gymnázia, kteří měli za úkol u prvního typu vět zhodnotit na škále jejich srozumitelnost, u druhého typu vět doplnit větná spojení. Tento krok jsem již realizovala se záměrem, že část získaných dat následně využiji ve statistické části své práce. Poté, co vedoucí experimentu zpracoval všechny připravené věty do podoby počítačového experimentu, zapojila jsem se do jeho administrace. Zúčastněné osoby měli za úkol číst předložené věty metodou self-paced reading. Získaná data opět zpracoval vedoucí experimentu a mě se následně dostala do rukou pouze ta data, která jsem potřebovala pro zodpovězení mnou položených hypotéz, které částečně vycházejí z **teoretické části** mé bakalářské práce. V té se věnuji za prvé některým teoriím, které se pokouší vysvětlit to, jak lidé s víceznačnými větami během čtení nakládají ve snaze o jejich interpretaci, dále jsem čtenáři poskytla exkurz mezi on-line metody, mezi něž patří i již zmíněná metoda self-paced reading, která byla v experimentu použita, a velice stručně jsem zmínila i samotný jev syntaktického primingu, na který se experiment sám o sobě primárně soustředil.

V **praktické části** jsem nejprve seznámila čtenáře se všemi kroky, které bylo nutno vykonat, abych získala potřebné soubory dat, které jsem následně statisticky zpracovala ve snaze zodpovědět na počátku položené otázky (**hypotézy**). Konkrétně

se jednalo o tvorbu experimentálních vět, sestavení a administraci dotazníků obsahujících tyto věty a následná administrace počítačového experimentu, který též obsahoval zmíněné věty, které měli účastníci za úkol číst metodou self-paced reading. Po detailním rozpracování všech zmíněných postupů následuje v rámci praktické části mé bakalářské práce kapitola věnující se **statistickému zpracování vybraných dat a následná interpretace zjištěných výsledků**.

V **hypotéze 1** jsem se věnovala otázce, zda je statisticky významný rozdíl mezi průměrnou dobou čtení vět víceznačných a jim příslušejících vět jednoznačných. Na dané téma jsem se podívala ze dvou úhlů. Za prvé (**varianta 1.A**) jsem se zaměřila na to, jaké průměrné doby čtení dosáhly jednotlivé osoby v celém souboru čtených vět víceznačných a v celém souboru čtených vět jednoznačných (průměrnou dobu čtení jsem zde tedy počítala u každého čtenáře zvlášť ze všech víceznačných a jednoznačných vět, které daná osoba v rámci experimentu přečetla). V druhém případě (**varianta 1.B**) jsem porovnávala průměrnou dobu čtení jednotlivých víceznačných vět a průměrnou dobu čtení jim příslušejících jednoznačných vět (průměrná doba čtení zde tedy byla spočítána z dob čtení dané věty v její víceznačné a jednoznačné alternativě u jednotlivých čtenářů). V obou případech jsem si potvrdila svůj předpoklad, že doba čtení vět víceznačných je statisticky významně delší než doba čtení vět jednoznačných. Víceznačné věty tedy čtenáři dle očekávání čtou pomaleji než věty jednoznačné.

Hypotézy 2 a 3 byly zaměřeny na faktory, které dobu čtení vět víceznačných mohou ovlivnit. V **hypotéze 2** jsem se ptala na to, zda existuje závislost mezi dobou čtení víceznačných vět a příslušnou hodnotou srozumitelnosti, která byla získána v dotazníkovém šetření. Zde jsem vycházela z předpokladu, že věty, které byly respondenty hodnoceny v dotaznících jako méně (hůře) srozumitelné, budou mít v rámci počítačového experimentu delší dobu čtení. Můj předpoklad se zde však nepotvrdil. Na základě korelační analýzy jsem zjistila, že mezi danými proměnnými není statisticky významná závislost. V **hypotéze 3** jsem připadala další proměnnou, kterou byl počet znaků ve větě. Ptala jsem se zde, zda je doba čtení víceznačných vět ovlivněna zároveň hodnotou srozumitelnosti a počtem znaků ve větě. V prvním kroku jsem spočítala korelační analýzu, kde jsem hledala jednoduchou závislost mezi

dobou čtení a počtem znaků. Zde jsem zjistila statisticky významnou mírnou pozitivní závislost. Pokud jsem však vzala v úvahu obě proměnné najednou jakožto prediktory doby čtení v rámci regresní analýzy, oproti očekávání se neprokázala statisticky významná možnost předpovídat dobu čtení pomocí počtu znaků ve větě a její hodnotě srozumitelnosti. To, že hypotéza 2 ani hypotéza 3 nebyly potvrzeny podle mých očekávání, mohla způsobit řada faktorů, z nichž některé jsem rozebrala v předchozí kapitole, tedy v **diskuzi**.

Slovníček pojmů

Pro upřesnění uvádím stručné vymezení některých lingvistických pojmů užitých v práci.

Valence (*Lopatková, 2008*) je jazykový jev, který hraje klíčovou roli při automatickém zpracování přirozených jazyků. Můžeme ji definovat jako schopnost slovesa vázat k sobě určitý počet syntakticky závislých jazykových jednotek (dalších slov). V rámci věty můžeme hovořit o doplněních valenčních, která tvoří tzv. aktanty neboli aktor, patient, adresát, původ a výsledek děje, a doplnění volná, tedy okolnosti děje jako jsou způsob, místo atd. Jednotlivé aktanty si však nejsou rovnocenné, nejsou pro porozumění větě stejně důležité – hovoříme o různé významové těsnosti valenčních pozic. Zde se dostáváme k tematice základové větné struktury (*Svozilová, 1997*) – tato struktura tvoří sémanticky a gramaticky dostačující, nedefektní větu, je tvořena slovesem a jeho obligatorními valenčními doplněními (*Lopatková, 2008*), tedy těmi, která není možné z věty vypustit, aniž by byl narušen její smysl případně gramatická správnost. Oproti tomu může věta obsahovat i doplnění fakultativní, kdy tyto valenční pozice není nutné ve větě obsadit, či doplnění potenciální (*Svozilová, 1997*), která můžeme z věty bez obav vypustit vyplývá-li jejich „obsah“ z kontextu či situace. Pokud bychom chtěli jít ještě hlouběji do problematiky, je zajímavé, že řídicí sloveso určuje morfematičnou podobu aktantů, tedy například to, v jakém jsou jednotlivá doplnění pádu (*Lopatková, 2008*).

Příručka *Odmaturuj z českého jazyka* (2002) podává následující definici jazykovědy s přehledem jejích hlavních disciplín (s. 5, 6):

Lingvistika neboli jazykověda je věda, která zkoumá a popisuje jazyk, včetně jeho užívání, a nejrůznější jazykové jevy. Můžeme hovořit o jazykovědě synchronní neboli popisné, která zkoumá jazyk v současnosti, a diachronní neboli historické, která se zabývá jazykem z hlediska jeho vývoje. Vnitřně se dělí podle toho, jakou složkou jazyka se zabývá. Základem je **gramatika** čili mluvnice, která popisuje

mluvnickou stavbu jazyka a dělí se na části jako hláskosloví, ortografie, morfologie (neboli tvarosloví, popisuje vnitřní stavbu slovních tvarů, slovní druhy a vysvětluje mluvnické významy slov ohebných) a syntax (čili skladba, zkoumá stavbu vět a souvětí), vedle gramatiky dále **lexikologie**, tedy nauka o slovní zásobě, která si všímá významů slov a vztahů mezi nimi, patří k ní jazykovědné disciplíny jako slovo tvorba, sémantika (nauka o významu slov), etymologie, frazeologie, onomastika a lexikografie a do třetice **stylistika** jakožto nauka o slohu.

Slovník spisovné češtiny pro školu a veřejnost (2001) a *Slovník cizích slov* (2002) uvádí k vybraným pojmům následující:

- ◆ **Lexikon:** slovník; **Lexikální:** týkající se slov, slovních spojení
- ◆ **Mluvnice, gramatika:** soubor zákonitostí, pravidel týkajících se soustavy slovních tvarů a ustálených způsobů tvoření vět, široce: stavba jazyka vůbec; nauka o mluvnické stavbě jazyka; dále např. *Sternberg* (2002, s. 600, 611) hovoří o **gramatice** jako o studiu jazyka na základě pravidelných vzorců vztahujících se k funkcím a vztahům slov ve větě
- ◆ **Morfologie:** nauka o tvarech, tvarosloví
- ◆ **Sémantika:** nauka o významu jazykových jednotek, významová stránka těchto jednotek
- ◆ **Syntax, skladba:** šířeji vztah, závislost mezi větnými členy, nauka o stavbě vět a souvětí, nauka o vzájemných vztazích skladebných prvků jazyka; jinak také rovina lingvistické analýzy, která se zaměřuje na vzorce, podle nichž uživatelé daného jazyka skládají slova v rovině věty (*Sternberg*, 2002)
- ◆ **Tranzitivnost:** přechodovost, přechodnost; **Tranzitivita:** v rámci logiky se jedná o druh vztahu mezi dvěma veličinami; **Tranzitivum:** přechodné sloveso, **Intranzitivum:** nepřechodné sloveso

V rámci své práce jsem pak dále operovala i s následujícími termíny, jejichž definici sem čerpala z *A dictionary of grammatical terms in linguistics* (Trask, 1993):

Tranzitivní sloveso je takové sloveso, které v rámci slovesné fráze (věty) vyžaduje tzv. **přímý předmět** (*direct object*), ten nejtypičtěji vyjadřuje tzv. *patient*,

což je sémantická role, která vyjadřuje entitu podstupující akci.

Intranzitivní sloveso je takové sloveso, které se v rámci slovesné fráze (věty) vyskytuje **bez přímého předmětu** (*direct object*).

Garden-path sentence je věta, která je konstruována tak, aby čtenáře během zpracování navedla k přiřazení nesprávné struktury. Zahrnuje tzv. **lokální víceznačnost** (*local ambiguity*), tedy víceznačnost, která je patrná jen když čtenář interpretuje část věty izolovaně a která následně zmizí, pokud je čtenář seznámen s celou větou.

Víceznačnost (*ambiguity*) je fenomén, ve kterém jednoduchý řetězec slov nese dva či více odlišných významů. Víceznačnost může být čistě lexikální, strukturální nebo kombinace obojího. Víceznačný řetězec jakožto **víceznačná věta** tedy koresponduje s dvěma či více odlišnými větami.

Literatura

DANCEY, Ch. P.; REIDY, J. *Statistics without maths for psychology : using SPSS for Windows*. 2. vydání. Prentice Hall, 2002. ISBN 0130336335.

EYSENCK, M. W.; KEANE, M. T. *Kognitivní psychologie*. Praha : Academia, 2008. ISBN 8020015590.

LOPATKOVÁ, M.; ŽABOKRTSKÝ, Z.; KETTNEROVÁ, V. *Valenční slovník českých sloves*. Praha : Karolinum, 2008. ISBN 978-80-246-1467-0.

MITCHELL, D. C. On-line methods in language processing: introduction and historical review. In CARREIRAS, M. a CLIFTON, C. J. (ed.). *The on-line study of sentence comprehension : Eyetracking, ERP, and beyond*. Psychology Press, 2004, s. 15-32.

Odmaturuj z českého jazyka. Brno : Didaktis, 2002. ISBN 80-86285-36-7.

RAYNER, K. Eye movements in reading and information processing : 20 years of research. *Psychological Bulletin*. 1998, roč. 124, č. 3, s. 372-422.

Slovník cizích slov. Praha : Encyklopedický dům, 2002. ISBN 80-90-1647-8-1.

Slovník spisovné češtiny pro školu a veřejnost. Praha : Academia, 2001. ISBN 80-200-0493-9.

SMOLÍK, F. Syntaktický priming porozumění větám v češtině. In ZELENKA, J.; MLS, K. (eds.). *Kognice 2008 (Sborník prací z konference)*. Hradec Králové : Gaudeamus, 2008, s. 85-92.

STERNBERG, R. J. *Kognitivní psychologie*. Praha : Portál 2002. ISBN 80-7178-376-5.

SVOZILOVÁ, N.; PROUZOVÁ, H.; JIRSOVÁ, A. *Slovesa pro praxi : valenční slovník nejčastějších českých sloves*. Praha : Academia, 1997. ISBN 80-200-0618-4.

ŠKALOUDOVOVÁ, A. *Statistika v pedagogickém a psychologickém výzkumu*. Praha : PedF UK, 1998. ISBN 80-86039-56-0.

TRASK, R. L. *A dictionary of grammatical terms in linguistics*. New York : Routledge, 1993. ISBN 0-415-08627-2.

Přílohy

Seznam příloh:

Příloha 1 – Věty A

Příloha 2 – Věty B

Příloha 3 – Věty doplněné do dotazníku A

Příloha 4 – Dotazník A

Příloha 5 – Dotazník B

Příloha 6 – Výsledky dotazníkového šetření, Dotazník A (1/2, 2/2)

Příloha 7 – Věty A použité v experimentu (1/2, 2/2)

Příloha 8 – Instrukce pro počítačový experiment

Příloha 9 – Tabulky: účastníci

Příloha 10 – Data k hypotéze 1, Varianta „čtenáři“

Příloha 11 – Data k hypotéze 1, Varianta „stimuly“

Příloha 12 – Data k hypotéze 2

Příloha 13 – Data k hypotéze 3

Příloha 1 – Věty A

Obecná struktura věty:

Podmět (*aktér*) + **přísudek** (*děj*) + **předmět** (3.p., *Komu?*) + **předmět** (4.p., *Co?*) + **příslovné určení** (předložkový tvar, *Kdy?*, *Kde?*, *Proč?*, *Jak?*)

Konkrétní věty:

Pytlák ukradl lvici lvíče ze křoví.
Záchranář zachránil zákaznici dítě při nákupu.
Školník natřel lavici nohy během prázdnin.
Pes pokousal důchodkyni nohu v parku.
Režisér vymluvil náhradnici roli před natáčením.
Muž obdivoval nápadnici kabelku při večeři.
Podezřelý udal žalobkyni důvody ze strachu.
Obžalovaný podplatil porotkyni kolegu před soudem.
Domácí rozmluvil nájemnici stížnost na chodbě.
Nepřítel zmrzačil důstojnici podřízeného během útoku.
Řezník zabil prasnici selata na dvoře.
Kamarád dohodil odhadkyni práci ve firmě.
Žhář podpálil stavení střechu před večerem.
Svědka přivedl policii podezřelého nad ránem.
Vrátný odvezl kolegyni tašku do bytu.
Dítě poprskalo výtvarnici obraz při exkurzi.
Selka přichystala slepici krmení do korýtka.
Host polil číšnici kalhoty v restauraci.
Brigádník vymaloval márnici sklad přes víkend.
Malíř namaloval účastnici podobiznu do novin.
Porotce přivezl vědkyni cenu na chatu.
Myslivec zastřelil bachyni kňoura v lese.
Asistent přitáhl hudebnici nástroj na pódium.
Plavčík resuscitoval plavčici manžela u bazénu.
Policista chytil uličnici papouška na ulici.
Soudce poslal trestankyni kytku do vězení.
Vyšetřovatel zatknul zrádkyni syna za krádež.
Pán odeslal hospodyni výplatu na konto.
Pomocník rozbil várnici víko v kuchyni.
Svalovec zbil zápasnici přítele u baru.
Mladík nabídl svědkyni úplatek v tajnosti.
Vtipálek ostříhal kadeřnici vlasy během spánku.
Nálezce dovezl ochránkyni ptáče do útulku.
Žalobce doporučil obhájkyni svědka před přelíčením.
Úředník ukázal poručnici děti v kanceláři.

Příloha 2 – Věty B

Obecná struktura slovního spojení/věty:

Podmět (aktér) + přísudek (děj)

Konkrétní slovní spojení:

Chlapec čte
Dědeček vypráví
Bratr snídá
Strýc obědvá
Tatínek večeří
Žák svačí
Doktor léčí
Novinář píše
Chirurg operuje
Teta šije
Babička plete
Sestra háčkuje
Maminka zašívá
Sousedka vyšívá
Pes žere
Kočka pije
Dítě jí
Osel žvýká
Štěně kouše
Kotě líže
Mladík zpívá
Žák recituje
Přátelé slaví
Fanoušci oslavují
Maminka zavařuje
Babička žehlí
Služka ždímá
Babička peče
Kuchařka vaří
Dítě mlsá
Mladík stopoval
Děvče maluje
Architekt kreslí
Žák počítá
Děvče zhaslo
Pán rozsvítil
Sousedka nakupuje
Pomocnice uklízí
Žena pere
Praktikant vyučuje

Příloha 3 – Věty doplněné do dotazníku A

Věty dodané konzultantem z Psychologického ústavu AV ČR, které byly přidány k větám A a dohromady vytvořily dotazník A.

Podmět + přísudek + předmět (3.p.) + dvouslovný přívlastek neshodný + předmět (4.p.)

Zpravodaj poslal redakci místních novin článek.
Nájemník donesl majiteli nového bytu peníze.
Úředník uložil vlastníkovu zanedbaného pozemku pokutu.
Taxikář doporučil trenérovi národního mužstva hotel.
Básník daroval manželce důležitého nakladatele kabelku.
Zlatník vnutil modelce obletované muži náramek.
Malíř ukázal řediteli velké galerie obraz.
Návrhář podal manekýně vyslané agenturou šaty.
Lékař nabídl pacientce trpící závratí lék.
Magnát půjčil provozovateli estrádních produkcí halu.
Dělník přidržel inspektorovi stavebního úřadu žebřík.
Kurýr předal zákaznici obchodního domu zboží.
Obchodník věnoval synovi významného dodavatele přehrávač.
Inspektor poskytl komisaři pověřenému případem spis.
Pomocník usnadnil znalci najatému advokátem práci.

Předmět (4.p.) + přísudek + přívlastek + podmět + příslovečné určení (předložkový tvar)

Ovci vyruší ukřičená vrána ze spaní.
Jelena uloví zákeřný pytlák v lese.
Orla klovne rozčilená straka do křídla.
Lišku postrčí unavený osel u vrat.
Vydru vytáhne vyhladovělý vlk z vody.
Psa ucítí vyděšená ryba na břehu.
Viníka prozradí zatčený komplic na policii.
Náčelníka zastoupí ctižádostivý plukovník při přehlídce.
Detektiva pochválil starý podvodník za mírnost.
Absolventa vybere mladá ředitelka v konkurzu.
Autora rozpozná nerudný kritik podle rukopisu.
Strážníka srazí mrštný útočník na zem.
Tajemníka vpustí pečlivý vrátný do objektu.
Dědečka namaluje malý chlapec na papír.
Konzultanta požádá bystrý právník o pomoc.
Výrobce prodal přepravci nebezpečných látek cisternu.

Příloha 4 – Dotazník A

Ukázka z jedné varianty dotazníku A.

DOTAZNÍK	<i>Verze I</i>
<p>Zadání: Přečtete si věty a u každé vyznačte na škále, nakolik se vám zdá daná věta srozumitelná, tedy jak snadno chápete, co tato věta říká. Hodnotit budete na škále od „velmi snadno srozumitelná“ (č.1) po „velmi obtížně srozumitelná“ (č.7).</p>	
<p>1) Tajemníka vpustí pečlivý vrátný do objektu.</p>	
velmi snadno srozumitelná (1) --- (2) --- (3) --- (4) --- (5) --- (6) --- (7) velmi obtížně srozumitelná	
<p>2) Pytlák ukradl lvici lvíče ze křoví.</p>	
velmi snadno srozumitelná (1) --- (2) --- (3) --- (4) --- (5) --- (6) --- (7) velmi obtížně srozumitelná	
<p>3) Zpravodaj poslal redakci místních novin článek.</p>	
velmi snadno srozumitelná (1) --- (2) --- (3) --- (4) --- (5) --- (6) --- (7) velmi obtížně srozumitelná	
<p>4) Záchranář zachránil zákaznici dítě při nákupu.</p>	
velmi snadno srozumitelná (1) --- (2) --- (3) --- (4) --- (5) --- (6) --- (7) velmi obtížně srozumitelná	
...	
<p>66) Úředník ukázal poručnici děti v kanceláři.</p>	
velmi snadno srozumitelná (1) --- (2) --- (3) --- (4) --- (5) --- (6) --- (7) velmi obtížně srozumitelná	
Pohlaví _____ Věk _____ Třída _____	
<u>DĚKUJEME ZA SPOLUPRÁCI !!!</u>	

Příloha 5 – Dotazník B

Ukázka z jedné varianty dotazníku B.

<u>DOTAZNÍK</u>	<i>Verze I</i>
Pohlaví _____	
Věk _____	
Třída _____	
<p>Zadání: Doplňte předložená slovní spojení libovolně tak, aby dávaly nově vytvořené věty smysl.</p>	
1) Chlapec čte	
2) Dědeček vypráví	
3) Bratr snídá	
4) Strýc obědvá	
5) Tatínek večerí	
6) Žák svačí	
...	
40) Praktikant vyučuje	
<u>DĚKUJEME ZA SPOLUPRÁCI !!!</u>	

Příloha 6 – Výsledky dotazníkového šetření, Dotazník A

1/2

Přehled hodnocení jednotlivých víceznačných vět (V0), které byly následně použity v počítačovém experimentu.

Tabulka, Hodnota srozumitelnosti												
Výsledky dotazníkového šetření, Dotazník A												
Stimulus	Průměrná hodnota srozumitelnosti V0			Minimální (min) hodnota srozumitelnosti V0			Maximální (max) hodnota srozumitelnosti V0			Směrodatní odchylka pro hodnotu srozumitelnosti V0		
	Muži	Ženy	Celkem	Muži	Ženy	Celkem	Muži	Ženy	Celkem	Muži	Ženy	Celkem
1	1,35	1,52	1,45	1	1	1	4	5	5	0,79	0,99	0,90
2	1,12	1,65	1,43	1	1	1	2	6	6	0,33	1,15	0,93
3	1,76	2,17	2,00	1	1	1	7	4	7	1,52	0,98	1,24
4	1,41	1,87	1,68	1	1	1	4	6	6	0,94	1,29	1,16
5	1,65	1,83	1,75	1	1	1	5	5	5	1,11	1,07	1,08
6	1,59	1,70	1,65	1	1	1	7	5	7	1,58	1,02	1,27
7	1,82	2,78	2,38	1	1	1	5	7	7	1,24	1,65	1,55
8	1,41	1,87	1,68	1	1	1	4	6	6	0,87	1,29	1,14
9	1,47	2,78	2,23	1	1	1	3	7	7	0,72	1,91	1,64
10	1,82	2,43	2,18	1	1	1	5	7	7	1,07	1,53	1,38
11	2,29	2,43	2,38	1	1	1	6	7	7	1,36	1,44	1,39
12	2,00	2,09	2,05	1	1	1	5	7	7	1,00	1,38	1,22
13	2,06	3,35	2,80	1	1	1	6	7	7	1,30	1,64	1,62
14	1,71	2,22	2,00	1	1	1	4	6	6	0,99	1,38	1,24
15	3,06	2,74	2,88	1	1	1	6	7	7	1,75	1,76	1,74
16	1,88	2,83	2,43	1	1	1	4	6	6	0,93	1,27	1,22
17	2,18	3,87	3,15	1	1	1	5	7	7	1,19	2,05	1,92
18	1,41	1,87	1,68	1	1	1	3	4	4	0,71	1,14	1,00
19	2,12	2,83	2,53	1	1	1	6	7	7	1,69	1,61	1,66
20	2,41	3,57	3,08	1	1	1	5	7	7	1,12	1,93	1,72
21	2,12	2,78	2,50	1	1	1	6	7	7	1,62	2,02	1,87
22	2,12	2,74	2,48	1	1	1	5	6	6	1,41	1,45	1,45
23	3,18	4,00	3,65	1	2	1	7	7	7	2,13	1,62	1,87
24	2,47	3,61	3,13	1	1	1	6	7	7	1,55	1,99	1,88
25	2,24	3,26	2,83	1	1	1	7	7	7	1,44	1,51	1,55
26	2,18	2,52	2,38	1	1	1	6	7	7	1,81	1,83	1,81
27	2,94	3,48	3,25	1	1	1	7	7	7	1,56	1,62	1,60
28	2,88	3,83	3,43	1	1	1	6	7	7	1,65	1,90	1,84
29	2,76	3,91	3,43	1	1	1	5	7	7	1,44	1,62	1,63
30	3,24	4,09	3,73	1	1	1	7	7	7	2,08	1,68	1,88
31	3,12	3,87	3,55	1	1	1	7	7	7	1,76	1,66	1,72
32	3,76	4,78	4,35	1	2	1	7	7	7	1,89	1,65	1,81

Tabulka hodnocení jednotlivých víceznačných vět (V0) jednotlivými osobami.

Tabulka se nachází v samostatném dokumentu s názvem Příloha 6, druhá část.

Příloha 7 – Věty A použité v experimentu

1/2

Víceznačné věty A (V0), které byly použity v počítačovém experimentu a jejich jednoznačné alternativy (V1) – 32 (dvojic) stimulů i s přiřazeným pořadím, se kterým se dané stimuly objevují ve výsledkových tabulkách.

Stimulus	Víceznačná věta (V0)	Jednoznačná věta (V1)
1	Vtipálek ostříhal kadeřnici vlasy během spánku.	Vtipálek ostříhal kadeřníkovi vlasy během spánku.
2	Selka přichystala slepici krmení do korýtká.	Selka přichystala kohoutovi krmení do korýtká.
3	Vrátný odvezl kolegyni tašku do bytu.	Vrátný odvezl kolegovi tašku do bytu.
4	Host polil číšnici kalhoty v restauraci.	Host polil číšníkovi kalhoty v restauraci.
5	Dítě poprskalo výtvarnici obraz při exkurzi.	Dítě poprskalo výtvarníkovi obraz při exkurzi.
6	Pán odeslal hospodyni výplatu na konto.	Pán odeslal sluhovi výplatu na konto.
7	Myslivec zastřelil bachyni kňoura v lese.	Myslivec zastřelil selatům kňoura v lese.
8	Pes pokousal důchodkyni nohu v parku.	Pes pokousal důchodci nohu v parku.
9	Soudce poslal trestankyni kytku do vězení.	Soudce poslal zpěvačce kytku do vězení.
10	Úředník ukázal poručnici děti v kanceláři.	Úředník ukázal poručníkovi děti v kanceláři.
11	Svěděk přivedl policii podezřelého nad ránem.	Svěděk přivedl policistům podezřelého nad ránem.
12	Malíř namaloval účastnici podobiznu do novin.	Malíř namaloval účastníkovi podobiznu do novin.
13	Nálezce dovezl ochránkyni ptáče do útulku.	Nálezce dovezl ochráncům ptáče do útulku.
14	Mladík nabídl svědkyni úplatek v tajnosti.	Mladík nabídl svědkovi úplatek v tajnosti.
15	Žhář podpálil stavení střechu před večerem.	Žhář podpálil domu střechu před večerem.
16	Domáci rozmluvil nájemnici stížnost na chodbě	Domáci rozmluvil nájemníkovi stížnost na chodbě.

Věty víceznačné (V0) a jejich jednoznačné alternativy (V1) – druhá polovina souboru.

Stimulus	Víceznačná věta (V0)	Jednoznačná věta (V1)
17	Porotce přivezl vědkyni cenu na chatu.	Porotce přivezl vědci cenu na chatu.
18	Pytlák ukradl lvici lvíče ze křoví.	Pytlák ukradl lvu lvíče ze křoví.
19	Řezník zabil prasnici selata na dvoře.	Řezník zabil prasnícím selata na dvoře.
20	Muž obdivoval nápadnici kabelku při večeři.	Muž obdivoval dívce kabelku při večeři.
21	Asistent přitáhl hudebnici nástroj na pódium.	Asistent přitáhl hudebníkovi nástroj na pódium.
22	Režisér vymluvil náhradnici roli před natáčením.	Režisér vymluvil náhradníkovi roli před natáčením.
23	Policista chytil uličnici papouška na ulici.	Policista chytil uličnickovi papouška na ulici.
24	Nepřítel zmrzačil důstojnici podřízeného během útoku.	Nepřítel zmrzačil důstojníkovi podřízeného během útoku.
25	Záchranář zachránil zákaznici dítě při nákupu.	Záchranář zachránil zákazníkovi dítě při nákupu.
26	Kamarád dohodil odhadkyni práci ve firmě.	Kamarád dohodil odhadci práci ve firmě.
27	Žalobce doporučil obhájkyni svědka před přelíčením.	Žalobce doporučil obhájci svědka před přelíčením.
28	Svalovec zbil zápasnici přítele u baru.	Svalovec zbil zápasníkovi přítele u baru.
29	Plavčík resuscitoval plavčici manžela u bazénu.	Plavčík resuscitoval ženě manžela u bazénu.
30	Vyšetřovatel zatknul zrádkyni syna za krádež.	Vyšetřovatel zatknul zrádci syna za krádež.
31	Podezřelý udal žalobkyni důvody ze strachu.	Podezřelý udal žalobci důvody ze strachu.
32	Obžalovaný podplatil porotkyni kolegu před soudem.	Obžalovaný podplatil porotci kolegu před soudem.

Příloha 8 – Instrukce pro počítačový experiment

Znění instrukce, tak jak byla zadána pokusným osobám při počítačovém experimentu (doplněna o názornou ukázkou postupu na první zácvičné položce).

„V průběhu experimentu uvidíte takhle 4 věty, Vaším úkolem je ty věty co nejrychleji přečíst. Jak jste si již zřejmě všiml/a, neuvidíte ty věty celé, jednotlivá slova budete odkrývat pomocí mezerníku (názorná ukáзка), vždy uvidíte jen jedno slovo. Když ty věty rychle přečtete (proklikání čtyřmi větami), zobrazí se Vám v následujícím okně otázka, která se vztahuje k jedné z těch předešlých 4 vět. Odpovíte na ni ANO/NE pomocí pravého či levého controlu (poukaz na nálepkou označené klávesy). Na odpověď máte zhruba 3 vteřiny, je tedy nutné se soustředit, pokud odpovědět nestihnete nebo odpovíte chybně, až tak moc se neděje, jen se pokuste dávat u další položky větší pozor, abyste těch chyb udělal/a co nejméně. Prvních 5 položek, tedy těch čtveřic, je zácvičných, pak plynule navazují položky experimentální, položek je něco málo přes 50 (záměrně uveden nižší počet), časově to tedy zabere zhruba 30 minut, záleží jak rychle budete číst. Je to poměrně dlouhé, ale pokuste se prosím vydržet až do konce a experiment nepřerušit. Já myslím, že to zvládnete, pokud nemáte otázky, půjdu tedy vedle, abyste měl/a klid. Na konec experimentu Vás upozorní nápis na monitoru, tak mě pak zavolejte.“

Příloha 9 – Tabulky: účastníci

V následujících tabulkách se nacházejí rozšířené údaje (věk a pohlaví) o účastnících dotazníkového šetření a počítačového experimentu.

Tabulka, Dotazník A						
Účastníci pilotního výzkumu, Dotazník A						
Věk	Třída septima		Třída 2.A		Celkem	
	žena	muž	žena	muž	žena	muž
16	-	-	4	2	4	2
17	4	1	8	2	12	3
18	7	11	-	-	7	11
19	-	1	-	-	-	1
Celkem	11	13	12	4	23	17
	24		16		40	

Tabulka, Dotazník B						
Účastníci pilotního výzkumu, Dotazník B						
Věk	Třída oktáva		Třída 2.B		Celkem	
	žena	muž	žena	muž	žena	muž
16	-	-	4	1	4	1
17	-	-	6	3	6	3
18	3	3	-	1	3	4
19	14	7	-	-	14	7
20	1	-	-	-	1	-
Celkem	18	10	10	5	28	15
	28		15		43	

Tabulka, Počítačový self-paced reading experiment																					
Účastníci počítačového experimentu																					
Pohlaví	Věk																				Celkem
	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	34	40	43	49	50	
Žena	1	-	-	1	1	5	3	-	3	-	3	1	1	1	1	1	1	1	1	1	26
Muž	-	1	1	2	-	2	1	1	1	1	-	1	-	-	-	-	-	-	1	-	12
Celkem	1	1	1	3	1	7	4	1	4	1	3	2	1	1	1	1	1	1	2	1	38

**Příloha 10 – Data k hypotéze 1,
Varianta „čtenáři“**

Data použitá pro výpočet t-testu, Varianta „čtenáři“.

Průměrná doba čtení jednotlivých subjektů					
Subjekt	Víceznačné věty (V0)	Jednoznačné věty (V1)	Subjekt	Víceznačné věty (V0)	Jednoznačné věty (V1)
1	4035,50	4047,65	31	2532,40	2549,01
2	5377,79	4380,72	32	5147,84	4319,79
3	7649,41	7189,79	33	3165,93	3065,88
4	3967,31	3834,72	34	5136,32	5026,01
5	2080,34	2173,32	35	5001,58	5208,34
6	2798,12	2545,83	36	4553,87	3753,02
7	2030,08	2013,63	37	3304,11	3591,80
8	3780,80	3813,16	38	2113,92	2042,99
9	1212,15	1080,27	39	2414,29	2407,75
10	2087,50	1982,36	40	3153,87	2991,25
11	1710,81	1908,92	42	3129,92	2779,89
12	2752,25	2451,95	43	2808,84	2836,15
13	4295,73	4510,87	44	2902,50	2824,43
14	4731,19	4645,25	45	3697,55	3055,09
15	2138,74	1868,61	46	1474,49	1420,56
16	3126,61	2967,73	47	2002,88	2041,74
17	3745,46	3500,71	48	2418,78	2209,08
18	2610,98	2391,97	49	2234,71	2290,31
19	5339,87	4885,86	50	1793,81	2080,77
20	4196,48	4045,63	51	2484,51	2718,04
21	3530,15	3754,94	52	4588,76	3583,13
22	4244,22	3935,19	53	3405,27	3508,46
23	2165,62	2229,98	54	2267,72	2298,57
24	2652,97	2668,09	55	2417,47	2474,17
25	4489,35	4468,40	56	1911,64	1704,04
26	3146,61	2814,49	57	2795,99	3095,28
27	3894,39	4068,15	58	3265,42	2987,49
28	3089,22	2889,17	59	4955,05	4159,12
29	2422,37	2610,48	60	2341,05	2180,76
30	3116,13	2817,67			

**Příloha 11 – Data k hypotéze 1,
Varianta „stimuly“**

Data použitá pro výpočet t-testu, Varianta „stimuly“.

Průměrná doba čtení jednotlivých vět		
Stimulus	Víceznačné věty (V0)	Jednoznačné věty (V1)
1	3052,23	3007,40
2	3168,00	3011,57
3	2894,98	2986,96
4	3144,54	2625,88
5	3809,02	3298,10
6	3389,19	2962,78
7	3230,01	3216,58
8	2916,46	2659,32
9	3068,81	2912,06
10	3468,62	3243,42
11	3352,38	3339,72
12	3373,60	3302,50
13	3289,56	2958,44
14	3190,44	2804,74
15	3163,57	2942,89
16	3328,89	2933,35
17	3106,12	3545,04
18	2957,22	2881,40
19	2952,22	3012,24
20	3312,67	2798,16
21	3498,93	3183,20
22	3435,83	3267,74
23	3358,44	3309,15
24	3519,48	3622,27
25	3078,57	3105,07
26	3333,42	3253,60
27	3378,34	3353,56
28	3500,49	3464,50
29	3377,76	2978,00
30	3269,44	3380,02
31	2956,62	3071,86
32	3166,07	3243,40

Příloha 12 – Data k hypotéze 2

Data použitá pro výpočet korelačního koeficientu a příslušné statistické významnosti.

Stimulus	Hodnota srozumitelnosti víceznačných vět (V0)	Průměrná doba čtení víceznačných vět (V0) v ms
1	1,45	3052,23
2	1,43	3168,00
3	2,00	2894,98
4	1,68	3144,54
5	1,75	3809,02
6	1,65	3389,19
7	2,38	3230,01
8	1,68	2916,46
9	2,23	3068,81
10	2,18	3468,62
11	2,38	3352,38
12	2,05	3373,60
13	2,80	3289,56
14	2,00	3190,44
15	2,88	3163,57
16	2,43	3328,89
17	3,15	3106,12
18	1,68	2957,22
19	2,53	2952,22
20	3,08	3312,67
21	2,50	3498,93
22	2,48	3435,83
23	3,65	3358,44
24	3,13	3519,48
25	2,83	3078,57
26	2,38	3333,42
27	3,25	3378,34
28	3,43	3500,49
29	3,43	3377,76
30	3,73	3269,44
31	3,55	2956,62
32	4,35	3166,07

Příloha 13 – Data k hypotéze 3

Data použitá pro výpočet korelačního koeficientu, vícenásobné lineární regrese a příslušných statistických významností.

Stimulus	Průměrná doba čtení víceznačných vět (V0) v ms	Počet znaků ve víceznačných větách (V0)	Hodnota srozumitelnosti víceznačných vět (V0)
1	3052,23	47	1,45
2	3168,00	44	1,43
3	2894,98	37	2,00
4	3144,54	40	1,68
5	3809,02	44	1,75
6	3389,19	39	1,65
7	3230,01	41	2,38
8	2916,46	37	1,68
9	3068,81	42	2,23
10	3468,62	42	2,18
11	3352,38	45	2,38
12	3373,60	45	2,05
13	3289,56	42	2,80
14	3190,44	42	2,00
15	3163,57	44	2,88
16	3328,89	46	2,43
17	3106,12	38	3,15
18	2957,22	35	1,68
19	2952,22	38	2,53
20	3312,67	43	3,08
21	3498,93	45	2,50
22	3435,83	48	2,48
23	3358,44	44	3,65
24	3519,48	53	3,13
25	3078,57	46	2,83
26	3333,42	41	2,38
27	3378,34	51	3,25
28	3500,49	39	3,43
29	3377,76	46	3,43
30	3269,44	45	3,73
31	2956,62	43	3,55
32	3166,07	50	4,35