	Detecting Errors in Corpus Annotation	Introduction	Detecting Errors in Corpus Annotation	Effects of Annotation Errors	Detecting Errors in Corpus Annotation
	Extractileurum University of Takingen		Extractileurum University of Takingon	<ul> <li>Less reliable training of NLP technology</li> </ul>	Extractileurs University of Takingen
Detecting Errors in Corpus Annotation	Ellants of Annualism Errors How to attain high-quality		Electric of American Errors Here to attain high quality	<ul> <li>van Halteren et al. (2001): a tagger trained on WSJ (Marrue et al. 1993) performe significantly worse than</li> </ul>	Barts of Armanian Errors Front to Arman Suph quality
On Variation Detection (http://decca.osu.edu)	Variation Antantion Computing validies regrams	Corpora with "gold standard" annotation are used	Part of Speech Variation datastion Computing variation regions	one trained on LOB (Johansson 1986)	Valid of Speech Valiation detailion Computing valiation rugsam
	Anaparate analytics Analytics the MLL Analytics where ballies	<ul> <li>for searching for linguistically relevant patterns</li> </ul>	Independent aufdersta from Vergraups angulation Annalis for the INELI Annalis for the INELI	<ul> <li>Less reliable evaluation of NLP technology</li> </ul>	Annaho to the MGJ
Detmar Meurers	Constituency Variation detection	Cost and the second to second to the form a second sub-model.	Constituency Veriates Assertion	<ul> <li>Van Halteren (2000): 13.6%–20.5% of cases where WPDV tagger disagrees with BNC-sampler annotation,</li> </ul>	Constituency Variation determine
University of Tubingen	WELFARE study Annale Description	markup process, which can include errors through	Will new work Annah	cause is error in BNC-sampler (0.3% error, Leech 1997). Error rates for other corpora much higher.	WELL near study Annale
	Computing nor registers Results for TEAR Increasing small	<ul> <li>automatic processes</li> </ul>	Computing our in-grams Results for TEXER Increasing small	<ul> <li>Padro &amp; Marquez (1998): because of errors in the testing data, cannot tell which of two tangers is better</li> </ul>	Computing our regrams Results for TADER Increasing senal
CLARA Thematic Training Course on Methods and Technologies for Consolidating and Harmonising Treebank Annotation	Resilts Dependency	<ul> <li>numan annotation or post-editing</li> </ul>	Reads Dependency	<ul> <li>Low precision and recall of queries for already rare</li> </ul>	Results Dependency
UFAL, Charles University, Prague December 13–16, 2010	Nature of departmention Indirect annotation Increasing small		Notice of dependencies Industry arminister Instant arminister	<ul> <li>Inguistic phenomena</li> <li>Meurers (2005): low precision of queries for verbal</li> </ul>	Nation of departmention Industrumentation Increasing small
	Summary		Summary	complex patterns since certain finite and non-finite verb forms are not reliably distinguished by German taggers	Summary
	YEVERULA'		YEVERULA'		"NURCEA"
Effects of Annotation Errors	Detecting Errors in Corpus Annotation	Effects of Annotation Errors	Detecting Errors in Corpus Annotation	How to obtain high quality annotation	Detecting Errors in Corpus Annotation
Searching for linguistic phenomena: The role of precision	University of Takingen	Searching for linguistic phenomena: The role of recall	University of Takingen		University of Takingen Introduction
<ul> <li>By precision of search we are referring to:</li> </ul>	Ren of American Errors Has to deale high quality Part of Speech	<ul> <li>By recall of search we are referring to:</li> </ul>	Ren of American Errors Has to deale high quality Part of Speech	<ul> <li>Annotate corpus independently several times, then test interannotator agreement (Brants &amp; Skut 1998: Artstein &amp;</li> </ul>	Part of American Trans New York and National Speech
Of the results to the query, now many represent the learner language patterns searched for?	Variation datastion Computing variables n.grams Independent acidence from	<ul> <li>How many of the intended examples that in principle are in the corpus are in fact found by the query?</li> </ul>	Variation datastion Computing variables n.grams Independent acidence from	Poesio 2009)	Variation detertion Computing variation in gram Independent acidence from
<ul> <li>False positives can result in two ways:</li> <li>Expression used in query also characterizes patterns</li> </ul>	Amate in the MGJ Annatation scheme leadlands	<ul> <li>Requirements on recall of search</li> </ul>	Amate in the MGJ Annatation scheme leadlands	the annotation agreement. Can the districtions made in the annotation scheme can be applied consistently	Annala for the NEJ Annala for the NEJ
other than the ones we are interested in.	Constituency Variation datastion Computing variation rugsame	<ul> <li>for qualitative analysis: Any results found useful, but danger of partial blindness where subcases are not</li> </ul>	Constituency Variation datastion Computing variation rugsame	Define adequate annotation scheme, with explicit	Constituency Variation Assertion Computing variation regions
<ul> <li>Requirements on precision of search</li> </ul>	WiLl near study Annalis Discontinuity	captured by query approximating target phenomenon.	WiLl near study Annalis Discontinuity	documentation and a list of problematic cases to achieve maximal agreement (Veutilainer & Javiner 1995;	Will name study Annalis Generationally
<ul> <li>for qualitative analysis: Needs to be high enough to find relevant examples among the false positives</li> </ul>	Results for TEXES Increasing small Results	<ul> <li>for quantitative analysis: Maximizing recall is crucial for reliable quantitative results.</li> </ul>	Results for TEXES Increasing small Results	Sampson & Babarczy 2003).	Results for TADER Increasing senal Results
<ul> <li>for quantitative analysis: For reliable results, very high</li> </ul>	Dependency Variation distantion	$\Rightarrow$ Where a query characterizing a target phenomenon is	Dependency Variation distantion	consistently identified and annotated uniquely	Dependency Variation distantion
language phenomena are concerned.	Industriantedation Increasing serial	expressed in terms of annotation, high annotation quality is important, and essential for quantitative analysis.	Industriantelation Increasing small	<ul> <li>appendix of difficult cases and how to resolve them crucial</li> </ul>	Industration Investig seal
<ul> <li>As known from zipris curse, most things occur rarely</li> </ul>					UNIVERSITAT
	4/21		S/M		6/91
Our research questions	Corpus Annotation	Variation Detection for POS Annotation	Corpus Annotation	Classifying variation	Corpus Annotation
<ul> <li>How about automatic methods for error detection?</li> <li>Detection can feed into repair as second stage of</li> </ul>	Introduction Electric American Error	(Biolinion & Wolfeld Ecology	Introduction Electric American Errors	The loss to also diving an inter the in the context.	Introduction Electric distribution Electric
correction (cf. also Oliva 2001; Blaheta 2002).	Ren to each top qualty Part of Speech Vacation docation	<ul> <li>POS tagging reduces the set of lexically possible tags</li> </ul>	Res to each typ quality Part of Speech	<ul> <li>The way to classifying variation lies in the context.</li> <li>The more similar the context of the occurrences, the</li> </ul>	Part of Speech
<ul> <li>What can be done for annotation of language in general?</li> <li>Can errors be found in common "rold standard" corpora</li> </ul>	Computing validities regrams Independent activities Improve ampliation	to the correct tag for a specific corpus occurrence.	Computing values in sparse Integration actions from Improve amplation	more likely the variation is an error. A simple way of making "similarity of context" concrete	Computing variation ruppers Independent autikense from Vergraupe angulation
regarding their	Annation where bashests Constituency	with more than one annotation.	Annation where bashests Constituency	is to say it consists of	American of the Mod American of the Mod
<ul> <li>syntactic annotation (Dickinson &amp; Meurers 2003b) Boyd,</li> </ul>	Variation detartion Computing variation regrams WELL new study	<ul> <li>Variation: material occurs multiple times in corpus with different annotations</li> </ul>	Variation detartion Computing variation regrams WELL new study	<ul> <li>which immediately surround the variation, and</li> <li>spanice identity of contents</li> </ul>	Variation detertion Computing variation ruppant WELL name study
discontinuous syntactic annotation (Dickinson & Meurers 2004)	Dameninuky Computing sar ingrams Parada for TSUER	<ul> <li>Variation can result from</li> </ul>	Diametinuky Computing sair nigrama Results for TSDER	Extract all n-grams containing at least one token that is	Descripting out repairs Results for TGDR
<ul> <li>oepenoency annotation (Boyd, Dickinson &amp; Meurers 2008) including spoken language corpora (Dickinson &amp; Meurers 2005a).</li> </ul>	homening sould Results	<ul> <li>genuine ambiguity</li> <li>inconsistent, erroneous tagging</li> </ul>	homening sould Results	annotated differently in another occurrence of that n-gram in the corpus.	Innuity small Results
<ul> <li>Detection of annotation errors through automatic analysis of comparable data recurring in the corpus</li> </ul>	Variation detaillion Native of departmention Indicast annotation	<ul> <li>How can one find such variation and decide whether it's</li> </ul>	Variation detailion Native of departmention Indicast annotation	<ul> <li>variation nucleus: recurring unit with different annotation</li> <li>variation n-gram; variation nucleus with identical context</li> </ul>	Variation Asturation Nature of Aspendention Indirect annulation
DECCA NSF project (http://decca.osu.edu)	Insuring soul Summary	an amoguity or error?	Insuring soul Summary		kreasing weat Summary
<ul> <li>Dickinson (2005)</li> </ul>	UNIVERSITAT		UNIVERSITAT		UNIVIESITAT STUDINGIN

Computing variation n-grams  • Example from WSJ: Variation 12-gram with off  (1) to ward fif a hotelia selecter attempt by two European  • once annotated as a reposition (N), and • once a particle (RP) • Note: Such a 12-gram contains two variation 11-grams • Note: Such a 12-grams contains two variation 11-grams • Or adculate variation n-grams based on variation n-grams to obtain an algorithm efficient enough for large contrains • colaculate variation or grams based on variation n-grams to obtain an algorithm efficient enough for large contrains • colaculate variation or grams based on variation n-grams to obtain an algorithm efficient enough for large contrains • colaculate variation or grams based on variation n-drams to obtain an algorithm efficient enough for large contrains • colaculate variation or grams based on variation n-drams to obtain an algorithm efficient enough for large contrains • colaculate variation regrams based on variation n-drams to obtain and algorithm efficient enough for large contrains • colaculate variation regrams • colaculate variation • colaculate variation • colaculate variation • colaculate variati	Checkperson and an an and an an and an and an and an an an and an	Computing variation <i>n</i> -grams Agonthm 1. Calculate the set of variation unigrams in the corpus and atore them. 2. Extend the <i>n</i> -grams by one word to either side. For each resulting (n + 1) gram 4. don't ince the set a variation in the vary the cocurrences are taged. 3. Repeat tage2 unit we reach an <i>n</i> for which no variation <i>n</i> -grams are in corpus. Running this algorithm on the Penn Treebark's version of the WSJ, retrieves variation <i>n</i> -grams up to length 224.	Internet int	Computing variation n-grams Example: WSJ in Penn Resbark 3 	Conception Conceptin Conception Conception Conception Conception Conception C
Heuristics for classifying variation I. The length of the context	Detecting Errors in Corpus Annotation Detect Neuros Unionaly of Takingen Introduction Errors & Annotation Takingen	Heuristics for classifying variation II. Distrust the fringe	Detecting Errors in Corpus Annotation Detections University of Talogon Introduction Error & Annotation Error	Why does the non-fringe heuristic work?	Detecting Enters Corpus Annotal Dense Neuron Unionity of Tange Introduction
Idea: The longer the <i>n</i> -gram, the more likely the variation is an error.	New to emain high quality Part of Speech Variation detertion Computing variable regions Magazinet automotion	Idea: Morphological and syntactic properties are governed locally. The further the variation nucleus is away from the edge of the <i>n</i> -gram, the more likely it is an error.	Real to disain high quality Part of Speech Variation detertion Computing variable regions Independent automations	<ul> <li>Non-image neuristic: one element of recurring context around a recurring nucleus is generally sufficient to determine that a variation in an annotation is erroneous.</li> </ul>	New to assume high quality Part of Speech Variation detection Computing variation reg Response to assess to
Example: In a variation 184-gram, the nucleus <u>lending</u> varies between adjective ( $\omega$ ) and common noun (NN). $\leftarrow$ lending $\rightarrow$	Analos to to ALL Analos to to ALL Constituency Variation attention Computing variation regions VELI new study Results	Example: A variation 37-gram with the nucleus joined occurring as first word: (3) a. John P. Karals	Analy in the second sec	<ul> <li>Is this an artifact of the WSJ annotation or is there independent motivation for such a general heuristic?</li> <li>Interestingly, recent research on language acquisition</li> </ul>	Reads to the MLL Annuals on the MLL Annuals on the MLL Constituency Variation detection Computing variation rig VEL rates study Reads.
109 identical words Julien 74 identical words Here, NN is the correct annotation of this <i>n</i> -gram.	Computing our registers Results for TABR Results for TABR Results	<ol> <li>Joint P. Karais has</li> <li>joined the Phoenix , Ariz. , law firm of Brown &amp; Bain . Mr. Karalis , 51 , will specialize in corporate law and international</li> </ol>	Geografic our regions Results for TABR Investig small Results	How do humans discover and learn categories of words?     His results show that humans seem to make use of	Computing our ruge Results for TABER Increasing small Results
Note: Heuristics independent of corpus, tagset, or language.	Summary	law at the 110-lawyer firm . Before joining Apple in 1986 , The context preceding the 37-gram shows: • In a. In ever must be tagged as past tense (vec), • in b. as past participle (ves).		exactly such non-fringe patterns ( <i>frames</i> ) to learn categories!	Summary UNIVERSITAT
Independent evidence from language acquisition	Detecting Errors in Corpus Annotation Detroit Moures University of Takingen	Independent evidence from language acquisition	Detecting Errors in Corpus Annotation Dense Mesons University of Takingen	Results for the WSJ	Detecting Errors Corpus Annotati Demer Movers University of Takings
<ul> <li>Mintz (2002) shows that lexical co-occurrence information of an element surrounded by a frame (i.e., XV) leads to categorization in adults.</li> <li>Mintz (2003): frequent frames supply robust category information, consistent aroses catilia language corpora.</li> <li>Example for a frame from CHILDES (MacWheney 2000): - you catil - you catil</li></ul>	Hencechon Research areas hang and an Petro of Species Merror and an antipart Merror and antipart Merror and antipart Merror and antipart Merror and antipart Merror and antipart Merror antipart Meror antipart Merror antipar	works most reliably when surrounded by a finane. The other same size contexts are much worse:	Electronics Part of Speech Part of Speech P	<ul> <li>Of the 2.485 distinct variation nuclei (types) 6 ≤ n ≤ 2.44:</li> <li>4.245 wereas (17.84%).</li> <li>Correcting the instances of these variation nucle by hard yields 417 bitem corrections.</li> <li>9.9 are genuine ambiguites</li> <li>9.1 are the first word 7 parts. 4 were 8 parts - relearce of hourists to prefer long context</li> <li>9.9 are prefer to ord the rogan. anying between the specific types. The first word of the rogan. anying between the 9.1 are the first word of the rogan. anying between the 9.1 of 7.141 distinct non-fringe variation n-gram types</li> <li>9.1 of 7.141 distinct non-fringe variation n-dist in the WS3, the method has a POS error tecal of at least 17%.</li> </ul>	Included of the second of the

Feedback for revising annotation scheme	Detecting Errors in Corpus Annotation	Related work on POS error detection	Detecting Errors in Corpus Annotation	Summary for POS error detection	Detecting Errors Corpus Annotati
For 140 of the 2436 erroneous variation nuclei, the variation was clearly incorrect, but which tag is the correct one is unclear from the guidelines (Statholm 1990). Example: Salomon <u>Brothers</u> Inc <u>Brothers</u> is tagged = 27 limes as proper noun (wer) = 22 as plural proper noun (wer) = 24 as plural proper noun (wer). = 24 as plural proper noun (wer). = 32 as plural proper noun (wer). = 32 as plural proper noun (wer). = 32 which may be a documented more explicitly or possibly eliminate, e.g.: = organ vs. common nours. = oretain types of noun-adjective homographs	Bandrahmann Hennessen	<ul> <li>Work with another focus, which could be combined with our consistency-checking approach.</li> <li>Denring and searching be lighting of Staging and the second second second second second second second mocristeterises are mostly possible bigman.</li> <li>Sparse Makov transdocers used to detect anomales, i.e., rare local tag patterns (Eskin 2000). → Inconsidencies are mostly recurst, not rates.</li> <li>Using passing bilanes to detect lishomed anotation of the second second second second second Using passing bilanes to the second second Using passing bilanes to the second second Using passing bilanes and the second second Using passing bilanes and the second second Using passing bilanes and second second Using passing bilanes and the second second second second second second second second relations 2000, Anny et al. (1997). Taggers detect consistence registant E.</li> </ul>	Benderhander Heinderhander Keiter Kei	<ul> <li>We discussed a detection methods for POS annotation errors in gold-standard corpora;</li> <li>detext variation within comparable contexts</li> <li>details uso variations are row ambiguity using general heuristics</li> <li>bids relies on multiple corpus occurrences of a particular word with different annotations – provide valuation for two received gold-standard valuation showed be method detacts errors in the WSJ with</li> <li>42.8% precision</li> <li>17% estimate recall</li> <li>Qualitative inspection of the detected variation can provide valuable feedback for cancelation scheme (rajdesign and documentation.</li> </ul>	Benefitier and the second seco
Variation Detection for Syntactic Annotation (Dokimon & Meurer 2008), 2004. Boyd, Dickmon & Meurer 2007) - Lerik tyr dapy viraition detection to the syntactic annotation in treebankal - Isovac nu boyntacticul annotated sentences be compared for tha? - Variation detection is closely related to interanonator agreement testing for multiply annotated corpus. - Isovac multiple annotation of the same sentences - compared to thai? - Closer (1972) nal Brants & Stu (1980) present agorithm for detecting aftherences in annotation. - agorithm is annotation-driven, asymmetric, and aminica-based = Wear les looking for a data-driven, symmetric, string-based approach.	Electrical Carteria in Carteri	Defining variation nuclei for syntactic annotation How can we obtain a data-driven definition of a variation nucleus as the unit of data on which the comparison of syntactic annotation can be based? <b>Problem:</b> No one to-ore mapping between word and label, as with part of speech. <b>Idea:</b> Decompose variation nucleus detection into series of nums for all relevant string lengths, more specifically a define one-to-one mapping between string of a given length and the label for that string • perform nums for strings from length 1 to longest constituent in corpus	Electrical Control of	Defining variation nuclei for syntactic annotation How to compare annotation for syntactic variation nuclei • To obtain a uniform mapping from strings to labels • assign all non-constituent occurrences of a string the special label ka. • Only compare categories assigned to the entire nucleus. • This intertionally grones the internal structure, • which is label nito account when shorter strings are cheeded.	Education Growth Annotate Service Service Ser
Examples from the WSJ corpus • variation between two syntactic category labels: • (a) maturity <u>metricategory</u> between as <u>WP</u> more • Variation between constituent and non constituent • using <u>weight weight </u>	<section-header></section-header>	Computing the variation nuclei of a treebank A simple way to calculate all variation nuclei: • so cal all stoches of length with category label or su. • at the all stoches of length with category label or su. • at the non-varying stretches <b>Problem:</b> Inefficient generate and refs method considering all stoches of strings starting at any position in the corpus. <b>Insight:</b> • The way we have set things up, variation involves at least one constituent occurrence of a nucleus. • Only strings analyzed as constituent scorewhere in corpus needs to be compared to annotation of other occurrences of that string.	<section-header><text><text><section-header></section-header></text></text></section-header>	Computing variation <i>n</i> -grams for a treebank Agonthm For each constituent length / (1 ≤ / ≤  longest-constituenti): 1. Compute the set of noclei: a) Find all constituents of length <i>i</i> : store them with their bind on the constituents of length <i>i</i> : store them with their add m, tore scath non-constituent of length <i>i</i> , add m, tore scath non-constituent occurrence 2. Compute variation nuclei set as: a all nuclei form set 1 with more than one takel 3. Generate variation <i>n</i> -grams for these variation nuclei, just as defined for part of speech annotation	Andread States and a second se

A case study: Applying the method to the WSJ • Two types of syntactic information in the PennTreebank3 (Marcus, Santonin, Macrioleanez & Taylor 1999); • estadia category generally determined by • estadia the covered ating and • syntactic function (skid) obtermined by • syntactic function (skid) obtermined by • syntactic function (skid) obtermined by • Beforus on the syntactic category. • TodErnRegistry (Laziur et al. 2002) converts e.g., transport for monour (Mar) register • veriation or grant test based on node labels only	<section-header><text><text><section-header></section-header></text></text></section-header>	Dealing with unary trees • unary branch causes same string to be annotated by two distinct categories • would be detended as unitation in annotation eliminate unary branches and relabel with mother/staughter category label, adding 70 new labels to original 27. • Example: NP OP OP DP DP DD 10 million		Constituent lengths in the WSJ $u_{ij} = \frac{1}{2} \int_{1}^{1} \int_{1}$	<text></text>
<ul> <li>Error detection results</li> <li>Total: 6277 district, non-fringe variation nuclei         district: each copus position is only taken into account         to longest variation rysamit accoust in         of longest context.</li> <li>We inspecting 100 randomly sampled examples:         <ul> <li>YTM STR 2000 (2011)</li>            &lt;</ul></li></ul>	Hendrig-Erection Hendrig-Erec	Misclassified Ambiguities I: Null elements • 10 of the 29 ambiguous nuclei in sample are null elements varying between two different categories. • WSJ anotacion inserted markets for arguments and adjuncts realized non-locally or unstated units of measurement (C. Bies et al. 1965, p. 69). • Example: TXP* (expetive) annotated as 5 or SBAR (3)# [Spars TXP*] may be inscaled (spars for the broke's action of the form of the moscie (spars) for the broke's action of the moscie (spars) for the broke'	Creating Groups and Annual States and Annual Sta	Misclassified Ambiguities I: Null elements Effect of eliminating variation detection for null elements • remove null elements from set of variation nuclei of length 1 • resulting number of non-tringe distinct variation nuclei 594 • 73, 39, 49, 50, 50, 50, 61, 72, 21, • elements are arrows • elements arrows	Entering Figure 2015 The second seco
Misclassified Ambiguities II: Coordination	Hendring-Erresting Corport-Provider Hendrick-Corport Hendrick-Corport Telescole Corport Hendrick-Corport Hen	Coordinate structure example interest in a flat coordinate structure: The amount covers large in the amount of t	Control Contro	<ul> <li>Related work on syntactic error detection</li> <li>CCGbank (Hockenmaier &amp; Steedman 2005): derived from Penn Trebank, King some errors: <ul> <li>e.g.: Under ADPR (Ho adverb has only one child, and it is tagged as NRP change it to RB:</li> <li>Blaheta (2002): discusses types of errors and some rules to dentify them</li> <li>e.g.: Tan N is occurring somewhere other than under a PP. It is likely to be a mistage.</li> </ul> </li> <li>Use &amp; Stimov (2004) search for unepected rules, using information about a node and its mother and daughter annotation can point to errors</li> </ul>	Checking Eric Checking Eric Heritage State Heritage

Summary for constituency error detection . We showed how one can extend the POS-error detection approach to syntactic annotation Blustrated with a case skulp abade on WSJ becknik that the method is successful (71% precision) in detecting inputation category annotation Approach supports the aspects of treebank improvement: . ensues I possible to a spects of treebank improvement extend to becknik for devicement of empirical descention to the deck of devicement of empirical descention extends to syntactic annotation, identifying desinctions difficult to maintain over entire corpus	Conception Annual Conception Concepting Concepting Concepting Concepting Concepting Conc	Discontinuous constituents - Discontinuous constituents (or equivalents) have been proposed in a vide range of syntacic Tamevorks, eg. : - Tee Adving Gamma - Categorial Gamma tows: toki reget with kom toki - Instruction-based Haad Dwine Phase Structure - Categorial Gamma (tows: toki reget advised) toki - Categorial Gamma (tows: toki reget advised) toki - Categorial Gamma (tows: toki reget advised) - Categorial Gamma (toki reget advised) - Categ	Execution Sector 2015 Terrest Control 2015 Terrest Contrest Control 2015 Terrest Control 2015 Terrest Cont	Some examples for discontinuous constituents  An English extraposition example:  (7) The man came into the nom who everybody loved  An English particle veb example:  (8) a. railed John pp.  Graman extraposition example (Brants et al. 2002):  (9) <u>Eff Mann Kommit, der lacht</u> a man comes, who laughs An and veb laughz comes:  Here Ein Mann der lacht is an NP constituent.	Compare Accession Accession of the accession of the acce
	TUNNGEN I		PRIVINGIA I		TUNNGEN 29/9
Treebanks and discontinuous constituents  • Treebanks which have been developed for languages with relatively free constituent order often represent to the constituent order often represent • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR Header State at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take a closer fork at: • FOR German we take at: • FOR GERMENT AT: •	Besteration of the second seco	The NEGRA/TIGER Treebanks	Execution Execut	An extraposition example (NEGRA corpus)	and the second s
Error detection for discontinuous constituents  • The variation n-gram method relies on the assumption that a continuous string can be mapped to a category. • Extend it to account for the fact that • the variation nuclei, an • here contexts are no longer required to be continuous strings, and • adapt the variation dassification heuristics accordingly.	Descriptions of the second sec	Adapting the algorithm to discontinuity Error detection for syntactic annotation is broken down into runs for all constituent lengths (1 5 / 5 / lengest-constituent); • Constituent size includes conty be tokens that are a part of the constituent, not possibly intervening material. 1. Compute the set of nuclei: a) Frod autocantituent of size / store them with their category label b) For each peop of sizing stored as constituent of length /, add each non-constituent of currence with more C. The variation nucleis at is the set of all nuclei with more	Checking Carson is Corporated Annotation Corporated Annotation The Corporation of the Corporation of the Corporation of the Corporation The Corporation of the Corporation Corporation of the Corporation of the Corporation Corporation of the Corporation of the Corporation Corporation of the Corporation of	Notes on Variation Nuclei Discontinuous non-constituent occurrences To find al strings teah match a constituent in the corpus, we need to take discontinuous strings into account. The strings to be found may not be constituents: (10) in discerem Punkt seine sigt Born and London nicht einig . not a agreed . "Born and London do not agree on this point." (11) in discerem Punkt seine sigt Born and London elfereichtlich	Defecting Errora in Corpus Annotation Server Management Interdeduction Interdeduction Part of Speech Varian along water region Research and the Constitution of the Research and Research an
	University of the second secon	than one label 3. Generate variation n-grams for these variation nuclei.	University and Statements Stateme	nicht argenet auf Bullin and London cleany not agreed. Here, sich einig is an AP in (10), but a discontinuous non-constituent (= NIL) in (11).	Summarian Manual Appendixion Manual Appendixion Manual Appendixion Manual Appendixion Manual Appendixion Summary UNIVERSITAT TURINGEN

Notes on Variation Nuclei Limiting the occurrence of non-constituents Constituents can overlap with non-constituent occurrences: (2) Ohne diese Augazben, so gie Weltbank, sein without these openess is the world bank are die Marchenie iden Kapital The secondary is the world bank without these expenses the people are dead capital - The string die Manschen occurs twice: - once as a continuous constituent - once as a continuous constituent - one as a constituent overlaps with a non-constituent string, ignore the non-constituent string.	Characterization of the second	Computing variation nuclei efficiently Use tries of storage Tsat: Frical algorating/stoortinuous strings that match a string courring as a constituent in the corpus. Determine a tracticable domain for the search: Syntactic annotation: only consider strings within a sentence. How to the search: I entertioner strenzte and strenzte the search I. Generate every (potentialy discontinuous) substring of a sentence (2 <sup>-7</sup> - classe sentence length r) 2. Test to see which once match a constituent. Incommental method using all are as guide: 1. Incommental method using all are as guide: 2. Incommental match every (potentialy discontinuous) substring of a sentence with a path in the time.	betrackoution Menoperation M	<ul> <li>Which contexts for discontinuous constituents</li> <li>Idea: the more similar the context, the more likely variation in the annotation of a nucleus is an error.</li> <li>Previously: expanded context to left and right</li> <li>Nour- also expand in internal context, i.e., material contained within span of discontinuous constituent but not part of constituent itself</li> <li>How to do It:</li> <li>Incrementally add context adjacent to the nucleus.</li> <li>Why? The most local context freips the most with disambiguition.</li> <li>Require surrounding context for every terminal element of the nucleus in order for it to be <i>non-fringe</i>.</li> </ul>	betrackoution Menoperation M
Results on the TIGER corpus The Setup • Used TIGER treebark (Brants et al. 2002), a German newspaper corpus with 712.332 tokens in 40.020 sentencos • Evaluation of whether a detected variation points to an error was carried out by George Smith and Robert Languer of the TIGER project.	Description of the second seco	Results on the TIGER corpus Baseline, without context: • Method detects 10.964 variation nuclei. • 13% pointed to at least one token error in sample of 100. (95% cont. Here: ''102 (64%) - 2148 (19.6%) are errors) Using word contexts (non-fringe nuclei): • Resulted in 500 abortest non-fringe variation nuclei, abortest non-fringe - rety solidy on non-fringe heuristic • 80% pointed to at least one token error in sample of 100. (95% cont. Herc: '381 (72.6%) - 439 (87.8%) are errors) • Precision comparable to regular syntactic annotation (71% in Dickinson & Meurers 2003b).	2019 Constantion	Increasing recall The variation <i>n</i> -gram method for detecting annotation errors Finds recurring data and compares analyses in different corpus instances Uses there donted as a heuristic to determine when analyses should be annotated identically we way to increase recall: We way to increase recall: Redefine variation mudels to extend the set of what counts as recurring data to which annotation is compared. Redefine context and heuristics to obtain more variation n-grams predicted to be errors.	
Approach explored Using part-of-speech nuclei to increase recall • To increase the number of errors found, relax the requirements of what constitutes comparable strings • Redefine variation nuclei: POS instead of words Example (WSL corpus, PennTreebank) tagest, 45 tags): (13) a. Boeing on Fiddy said 0 a received by and the index of the strings in the form to a movie of the string words (POS tags) may regalitively impact the precision of error detection. • To use a more general recuting units (POS tags) may regalitively impact the precision of error detection. • To use a more general recuting units (POS tags) may regalitively impact the precision of error detection.	Control Contro	Approach explored           Identifying reliable contexts to maintain high precision           Example liutuating problem that shortest non-ringe heuristic does not ensure sufficient context:           (14) a. cripped* by a biter, dicade-long site that "T" togat_grided" by a biter, dicade-long site that the biter of the site that the site that the biter of the site that biter of the site that the biter of the site that biter of the site that biter of the site that biter of the biter of the site that biter o	Centring Constant Constant and	Heuristic 1: Shared complete bracketing Target 1: Variation with bracketing agreement, i.e., between nuclei wich are constituents (XP vs. YP) Both annotations agree on the bracketing Significantly more likely that variation in label is an error Exi: RB, JJ varies between NP and wrong ADJP in 4-gram: (15) a. This was <u>how Both Bracketing</u> to a dearup <i>al an ol-negolithy plant</i> . Heuristic 1: Shared complete bracketing is comparable context	Centring Constant Constant Annual Sector Constant Constant Constant Sector Constant

Heuristic 2: Shared partial bracketing	Detecting Errors in Corpus Annotation	Heuristic 2: Shared partial bracketing	Detecting Errors in Corpus Annotation	Heuristic 3: Shared vertical context	Detecting Errors i Corpus Annotatio
Target 2: Variation between constituent and non-constituent	Extractilizers University of Takingen Introduction Black of Annulation Errors Tenson and Annulation Errors	Heuristic 2: Require one extra word on side(s) without shared bracket.	Estima Macana Unitarialy of Takingan Introduction Electric of Annulation Errors New Josephic Macanation	Target 3: Variation in bracketing of nucleus, but shared bracketing in <i>n</i> -gram	Extract Mercers Unteraity of Takingen Introduction Effects of Annatolies Error
<ul> <li>a. crippled " by a bilter, decade-long strike that "T" began [pp [m]NI [1967/CD] and cut circulation in half</li> <li>b. its problems began [pp [m]NI [pp [1987/CD] and early 1988] when Te</li> </ul>	Part of Speech Variate Arterion Computing validate rupans Independent autores tem Inngelege amphilities Analis to the INLI Annalis to the INLI	Example: Erroneous variation for the variation nucleus VBG JJ NNS (17) he stayed inside the Capitol (17) he stayed inside the Capitol	Part of Speech Varian advances Computing variables rugsam hispansiset autores tem impage angulation Results for the NEL Annotation where busilesit	Example: erroneous variation for nucleus RB JJR IN CD: (19) a. will be diluted * to [net [or latightly/RB] less/JJR [than/IN_50/CD] %] after	Part of Speech Variation distantion Computing variation rules Integrating angulation Results for the INU Annualis for the INU
Logismute attachment difference here because     'n 1987 forms a complex VP with began, but     'n 1987 forms a complex VP with began, but     one word of surrounding context is not sufficient to     disfuguable.     Can we define a heuristic to reduce risk of attachment     ambiguites?	Continuery Verain and the open State and the open State and the open State and the State State and the State and the State and the State State and the State and the State and the State State and the State and the State and the State and the State State and the State and	Instant of thing to San Francisco]      Instant of the first source of the fi	Constancy Version available Constances of the Constances of the Neuronal Constances Neuronal Constances Ne	<ul> <li>be at all b</li></ul>	Constituency Varient waters no Service of the service Will near the Will near the Will near the Parties of the
Results for POS nuclei After generalizing the nuclei from words to POS, we obtain 50.369 variation nuclei for the VKJ 6.568 of which remain after removing nuclei which are single null elements (cf. Dichards Meurers 2004) 7. Gauge performance of POS nuclei 8. Sampled to Closes from 16.598 to examine by hand 8. Sampled to Closes from 16.598 to examine by hand 9.2% point to an error 9. 4.677 estimated cases of errors, which is a significant improvement in recall over 2.745 for word nuclei	Checkeleg Cherse in Carpor Areas Summer Cherse House Cherse Here C	Results with heuristics           • Heuristics select 6.34 granulation nuclei of 16.598:           • Heuristic 1 (Shared complete branketing): 1.337           • Heuristic 2 (Shared partial bracketing): 2.731           • Heuristic 2 (Shared partial bracketing): 2.731           • Heuristic 3 (Shared verifica context): 1.273           • Inspected random sample of 100 cases to judge precision:           • Meristic 1 61%           • Heuristic 1 61%           • Heuristic 2 (Shared verifica complexity)           • Estimate 4.337 entros tom 6.343 cases, 59% increase in recail on verificated 2.748 precision           • New heuristics cover most cases, approaching high precision of word nucleis increasing mecall.	Checking Christing Carport Andread Hermiter States International States of Annotation States	Limitations of POS nuclei Generalizing from word to POS nuclei is not always successful, i.e., POS class not fine grained enough. Example: variation trigram "tremains Jufor" (20) a virus that "T (see remains Jufor" (20) a virus that the see	Electrical Events Carpanet Annabel Series Lever Hereits and Annabel Hereits and Annabe
Alternatives ways to increase recall  • Use more general types of context (e.g., POS tags, Dicknoor 2005; Dicknoor & Maurers 2005b) • 8.715 shortest non-Hinge variation noted, with an • Could be contineed with the POS nucleus approach using the new heuristics. • Immediate dominance variation method (Dickinson & Meurers 2005c) based on RHSs of treebank rules • Overlage with shared complete bracketing cases when • RHS is complete sequence of POS tags • Used and not bracketing errors • babling and not bracketing errors • Separate slides on exploring endocentricity		Summary for increasing recall of constituency error detection • increased error detection recall for syntactic annotation by generalizing nature of comparable recurring unit • Generalized variation negle in variation ergam method to POS tags instead of using identical surface forms • Determined additional contextual heuristics for errors	Control Contro	Variation Detection for Dependency Annotation (Boyd, Dickinson & Meurers 2008) • A range of high-quality dependency treebanks for a variety of different languages are available, e.g.: • Prague Dependency Teebank (PDT) of Czech (Higle et al. 2001) • Alpino Dependency Teebank (PDT) of Czech (Higle et al. 2001) • Alpino Dependency Teebank (PDT) • Tabate for the state of the state • Tabate for the state of the state • Alpino Dependency paraiting highlighted by 2006 CoNLL-X Shared Task • Als fars are are aware, life work has been done on automatically detecting errors in dependency treebanks.	Construction Const

Dependency annotation Some characteristics • Dependency annotation • convention of the follow between words • convention of the follow between words • may include non-projectivity, i.e. dependency area can oro • Example from Tabanken05 corpus (Nivre et al. 2006): • <u>or sol</u> <u>or sol</u> (21) Deras ubbidning tar 345 daya Their education takes 345 daya		Corpora used for dependency error detection  • Explore approach on the basis of three devise dependency annotation schemes for three languages:  • Tabariendo corput of Sweddah (Nvr et al. 2008) • approx. 200.000 hims • approx. 200.0000 hims • approx. 200.000 hims • approx. 2	<text></text>	Adapting the method to dependency annotation 4. What is involved in applying the variation or gams method to dependency annotation? 4. We encode the head information rule of size 2. 4. We encode the head information rule of size 2. 4. We encode the head information rule of size 2. 4. We encode the head information rule of size 2. 4. We encode the head information rule of size 2. 4. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 2. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode the head information rule of size 3. 5. We encode thead information rule of size 3. 5. We encode thead informa	Conception of the section of the sec
Applying the variation <i>n</i> -gram method With the dependency annotated data encoded in this way, there are three different possibilities for errors: Errors in tabeling: SUBJ var, OGJ Errors in what the head is: OBJ vs. OBJ+R Tors in dependency identification: OBJ vs. NUL What needs to be added to the basic picture? Take the nature of dependency annotation into account in a defining the set of variations that need to be considered determining a notion of context sufficient to identify the variations which are errors — heuristics	Characterization of the second	Errors in dependency identification  • The existence or absence of a dependency is captured by variation with the special late NIL. • E.g. DT vs. NIL in the following Tabanken05 example: (2) • • • • • • • • • • • • • • • • • • •	Checkson States State	Accounting for the nature of dependencies  • Variation r-gram approach for constituency useful starting point, but how do we adapt it to the nature of dependencies?  • There are formal and injusistic issues in adapting the method from constituency to dependency annotation, including: • Overlap • Contiguity	<section-header><text><text><section-header><section-header><section-header></section-header></section-header></section-header></text></text></section-header>
Constituency and Dependency Overap Provide A state any of their daughters, unless one is properly included in the other. In a dependency representation, the same head may adependency pairs loweride in one token. Example: tri is the head of two dependencies in (21) Example: tri is the head of two dependencies in (21) Control of the other ot	Conception of the section of the sec	Constituency and Dependency Dealing with overlap (2) She showed the department chair the beauful old chair . ⇒ Set up variation detection to compare sets of all dependencies between the work in the nucleus. • e.g., for the case above: < showed, chair, (DO, IO) >	Conception of the second secon	Constituency and Dependency Contiguity • Which in traditional constituency frameworks, the sisters in a local tree are configuous, i.e., their terminal yield is a continuous string. • For dependency annotation, a dependency graph will other relate non-configuous elements. (4) Continuous Control of the control of t	Construction of the second sec



Heuristic 1: NIL internal context heuristic Example for case predicted to be an error (31) a. * Writechaftspohitik kill auf sich warten * economic policy lets on itself wat cer of our occur. Der Writschaftspolitik kill auf sich warten te conomic policy is a long time conting.	<section-header>because and a second s</section-header>	Heuristic 1: NIL internal context heuristic Example for case predicted not to be an error MO BED of the States (22) a. in det Vereinigten States in det vereinigten States but of the United States in den vergangenen zehn Jahren an der Vereinigten S in be paat ten years to be United States	OBJ itaaten States	Heuristic 2: Dependency context heuristic  • If the head of a variation nucleus is being used in the same function in all instances, the variation in the labeling of the nucleus is more likely to be an error. • Conversely, when the head is used differently, it is more likely a genuine ambiguity.	Beneric and a series of the se
Heuristic 2: Dependency context heuristic Example: head with two different functions $\rightarrow$ not an error (3) a. <i>i</i> den ena effer and a kormen in the one or other form b. <i>i</i> den ena effer and a kormen in den one or other different b. <i>i</i> den ena effer bada fladitioningama in the one or other directions	Betecting Errors is Corpus Anotation Insure Anotation Insure Anotation Insure Anotation Insure Anotation Insure Anotation Insure Anotation Insure Insure Anotation Insure Insu	Results Talbanken 05 • 197,123 tokens in 11,431 sentences in sections P and G • 210 different variation nuclei using non-finge heuristic • 29.95, precision (195 error nuclei) (thanks to Jaakim Nirve, Mattias Nisson and Eva Petersson for the evaluation • 274 error instances: • 136 datendersry identification • 139 dependency identification • observations: • common problems with advefabilit (70) than arguments (31)	Detecting Errors in Corpus Another Management Histocheckics Histocheckic	Results PDT 2.0         38,482 sentences (670,544 tokens) in full/amw section           553 different variation nuclei using non-fringe heuristic         480 cases after removing entres involving punctuation           - 480 cases after removing entres involving punctuation (indirect lands after recoding indirect Aure) and Aux C deputed (transits to Linka Hana and Jan Böghank for evaluation)           - 513 clashing contuined (60.5%)         69 dependency determination observation: genotifies due to after indirect annotation - scheme decision (coordination)           - componentiem Vindentiation         -	Detecting English Detecting English Detecting English Detecting
	Summary UNIVERSITAT TUNINGEN		Summary UNIVERSITAT TUNNGEN	adverbial of predicate vs. attribute of lower node	Summary UNIVERSITAT TUNINGEN
Results           Tiger0B           Only used sentences with lexically-rooted dependency structures, ignoring abstract and sublecial nodes.           1.575 retraction roucle, NLL internal context heuristic           4.815 protection (133 error nucle)           1494 error instances           • 61 bigenderschild (193 error nucle)           • 62 bigenderschild (193 error nucle)           • 63 bigenderschild (193 error nucle)           • 05 begenderschild (193 error nucle)           • 05 begenderschild (193 error nucle)           • 05 bigenderschild (193 error nucle)           • 05 bigenderschild (193 error nucle)           • 05 bigenderschild (193 error nucle)           • propositional argument vs. modifier distinction difficult           • 16 proteilschild (193 error nucle)           • 16 proteilschild (193 error nucle)           •	Helicity Const.	Outlook: Increasing recall           The issue         • word word dependencies are highly specific.           • How can they be generalized to increase the number of recurring dependency pairs limiting the recall of the variation detaction method?           Specific lexical properties of head important, e.g.:           Lexical information is known to improve PCFGs through head-lexicalization (e.g., Collins 1996)           To characterize the dependent, POS class may be sufficient (cl. subcategorization frame in lexicalization frame in lexicalization dependent, PUS class may be sufficient (cl. subcategorization frame in lexicalization theorem is expendent).           = Generalize from word-word to word-POS dependencies           • For nucle of an antodata as dependent, full, use head-dependent orientation of string we compare it.to.	Control Contro	Outlook: Increasing recall Tagset dependency • Use of word-POS dependencies is dependent on the granularity of the POS tagset used. • Tatawine Groups tas 40 coarse grained POS tags • PDT 2.0 desinguistance 4200 POS tags (Heig2 004) • For positional lagests, one can decide which positions of the tagset to use, o.g. • including case information takey to increase precision • distinguisting comparative and supertative adjectives could decrease recall	Helicity Error 1 Core Andrew Core Andrew Theorem Core Helicity Core Andrew Cor

Summary	Detecting Errors in Corpus Annotation	References	Detecting Errors in Corpus Annotation	Blevins, J. (1990). Syntactic Complexity: Evidence for Discontinuity and Multidomination. Ph.D. thesis, University of Massachusetts, Amherst, MA. Bonami, D. D. Godord, S. J.M. Mazandrin (1990). Constituency and word certer in	Detecting Errors in Corpus Annotation
We motivated the need for error detection in annotated	University of Takingen	Abailó & (ad.) (2002). Teachaster: Building and using sustantically associated	University of Takingen	French subject inversion. In G. Bourna, E. W. Hinrichs, GJ. M. Kruijff & R. T.	University of Takingen
corpore, and introduced the variation a-gram approach	Introduction Electric American Error	corpora. Dordrecht: Kluwer Academic Publishers.	Introduction Electric discussion Errors	Oehrie (eds.), Constraints and Resources in Natural Language Syntax and	Introduction Electric Acceleration
as an automatic error detection method	Here is strain high-quality	http://treebank.linguist.jussieu.fr/toc.html.	How to strain high quality	Semantics, Stanford, CA: CSLI Publications, Studies in Constraint-Based Lexinalism on 21-40	Here to strain high-quality
as an automatic error detection metriod.	Part of Speech	Abney, S., R. E. Schapire & Y. Singer (1999). Boosting Applied to Tagging and PP	Part of Speech	Boyd, A., M. Dickinson & D. Meurers (2007). Increasing the Recall of Corpus	Part of Speech
<ul> <li>Research on category learning in humans provides</li> </ul>	Computing variables in gramm	Attachment. In P. Fung & J. Zhou (eds.), Proceedings of the 1999 Joint	Computing variables in gramm	Annotation Error Detection. In Proceedings of the Sixth Workshop on	Computing variables in gramm
independent evidence for the notion of context used.	Impage amphilies	and Very Large Corners on 38-45	Respondent ausbaltum Gesche Angeleiten	Treebanks and Linguistic Theories (TLT-07). Bergen, Norway. URL	impage angulation
<ul> <li>The method successfully detects errors in</li> </ul>	Annual or scheme loadback	Anrawal B & B Srikant (1994) Fast Algorithms for Mining Association Bulas in	Annatation scheme leasthants	Bowl & M Dinkinson & D Meurers (2008). On Detertion Errors in Dependency	Annatation scheme loathants
<ul> <li>The method successibility detects errors in</li> </ul>	Constituency	Large Databases. In J. B. Bocca, M. Jarke & C. Zaniolo (eds.), VLDB 1994.	Constituency	Treebanks. Research on Language and Computation 6(2), 113–137. URL	Constituency
<ul> <li>part of speech</li> </ul>	Computing variation in gramm	Morgan Kaufmann, pp. 487–499.	Computing variation regiants	http://purl.org/dm/papers/boyd-et-al-08.html.	Computing variation in gramm
<ul> <li>constituency,</li> </ul>	Will near study Annals	Artstein, R. & M. Poesio (2009). Survey Article: Inter-Coder Agreement for	WU new study Annula	Brants, S., S. Dipper, S. Hansen, W. Lezius & G. Smith (2002). The TIGEH Tendhork. In Researchings of the Workshop on Tenchorks and Linguistic	Will new study Results
<ul> <li>discontinuous constituency,</li> <li>and dependency expectation</li> </ul>	Description of the second seco	Computational Linguistics. Computational Linguistics 34(4), 1–42. UHL http://www.mitoressiournals.om/doi/abs/10.1182/ooli.07.034.B2	Description of the second seco	Theories. Sozopol, Bulgaria. www.bultreebank.org/proceedings/paper03.pdf.	Description of the second second
<ul> <li>and dependency annotation</li> </ul>	Penalty for TADER	Birk F. (2003). Arboretum a Hybrid Traebank for Danish. In Proceedings of TLT	Results for TADER	Brants, T. & W. Skut (1998). Automation of Treebank Annotation. In Proceedings of	Results for TAXAN
<ul> <li>We showed that the method can provide significant</li> </ul>	Paralite	2003 (2nd Workshop on Treebanks and Linguistic Theory. Vilxjö, Sweden, pp.	Penalte	New Methods in Language Processing (NeMLaP-98). Syndey. http://www.coll.upl.sb.do/=thorstop/w/biostione/Broots_Skut NeMI aB98 os.co.	Paralto
feedback on annotation scheme distinctions which	Dependency	9-20.	Dependency	Bröker, N. (1998). Separating Surface Order and Syntactic Relations in a	Dependency
<ul> <li>are not sufficiently documented,</li> </ul>	Nature of Aspendermian	Bies, A., M. Ferguson, K. Katz & R. MacIntyre (1995). Bracketing Guidelines for	Nature of dependencies	Dependency Grammar. In Proceedings of the 17th International Conference	Notice of Approximities
<ul> <li>rely on representational choices not locally motivated,</li> </ul>	increasing small	for it's unero eduinubiteebank/roject. Oniversity of Pennsylvania.	Pressing scal	on Computational Linguistics (COLING) and the 36th Annual meeting of the	Pressing scal
<ul> <li>or cannot reliably be made based on the evidence found</li> </ul>	Summary	Blabeta D (2002) Handling price training and testing data. In Proceedings of the	Summary	Calder J (1997) On aligning trees in Proceedings of the Second Conference of	Summary
in the corpus,	UNIVERSITAT	7th conference on Empirical Methods in Natural Language Processing. pp.	UNIVERSITAT	Empirical Methods in Natural Language Processing. Brown University.	UNIVERSITAT
	TONNGEN 📥	111–116. http://www.cs.brown.edu/~dpb/papers/dpb-emnlp02.html.	"TONNGEN"	http://xxx.lanl.gov/abs/cmp-lg/9707016.	"TONNGEN"
	90.795		91/91		90795
chemia, E., I. H. Mintz, S. Bernal & A. Christophe (2009). Categorizing words union "trouvert frames": what cares linewidtle applying munoi about	Detecting Errors in Corpus Apartment	Diokinson, M. & W. D. Meurers (2005b). Detecting Errors in Discontinuous Structural Apartation. In Proceedings of the 43rd Apartol Meeting of the	Detecting Errors in Corpus Apartmice	Hajic, J., A. Bonmova, E. Hajicova & B. Vidova-Hiadka (2003). The Prague Dependency Technols: A Three Level Association Scenario. In Abelia (2002).	Detecting Errors in Corpus Apartment
distributional acquisition strategies. Developmental Science 12(3). URL	Extract Meuron	Association for Computational Linguistics (ACL'05), pp. 322–329.	Estimat Messare	chap. 7, pp. 103–127. URL	Datmar Meurana
http://dx.doi.org/10.1111/j.1467-7687.2009.00825.x.	University of Takingen	http://www.aclweb.org/anthology/P/P05/P05-1040.	University of Takingen	http://ufal.mfl.cuni.cz/pdt2.0/publications/HajicHajicovaAl2000.pdf.	University of Takingen
Collins, M. J. (1996). A New Statistical Parser Based on Bigram Lexical	Elate of American Errors	Dickinson, M. & W. D. Meurers (2005c). Prune Diseased Branches to Get Healthy Troop How to Elect Erropsource and Troop in a Troophark and Mitry & Mattern, In	Ellaris el Arrolador Erros	http://treebank.inguist.jussieu.tr/toc.ntml. Holiči I. P. Hardké & P. Polos (2001). The Branue Dependency Technols:	Ellaris el Arrolador Erros
Dependencies. In A. Joshi & M. Haimer (eds.), Proceedings of the Thirty: Enurth Annual Menting of the Association for Computational Linguistics	Here is all all high spacing	Proceedings of the Fourth Workshop on Treebanks and Linguistic Theories	How to strain high-quality	Annotation Structure and Support. In IRCS Workshop on Linguistic Databases.	Here is all all high quality
San Francisco: Morgan Kaufmann Publishers, pp. 184–191. URL	Part of Speech Velation detertion	(TLT 2005). Barcelona, Spain. URL	Part of Speech Validity deterior	Hepple, M. (1994). Discontinuity and the Lambek Calculus. In Proceedings of the	Part of Speech Values detertion
citeseer.nj.nec.com/collins96new.html.	Computing variation in gramma Independent authors from	http://ling.osu.edu/~dm/papers/dickinson-meurers-ttt05.html.	Computing variation regiants Independent acidence from	15th Conterence on Computational Linguistics (COLING-94). Kyoto.	Computing variation regiants Independent acidence from
Dickinson, M. (2005). Error detection and correction in annotated corpora. Ph.D.	tempage amphilien Anada to the INL	Conference on HPSG. Edinburgh: University of Edinburgh. on 101-106	language angulation Angula to the ML	ttp://ttp.dcs.snet.ac.uk/nome/neppie/papers/coiings4.ps.	language angulation Annula for the INLU
thesis, The Unio State University.	Annatation scheme loadback	http://www-csli.stanford.edu/~sag/papers/warlpiri.ps.	Annatation scheme loatilacti	Tagger Using a Reliable Parser and Plain Text Corpora. In Proceedings of the	Annatation scheme loathants
Annotation In Proceedings of the 10th Conference of the Evaneer Chanter of	Constituency	Dowty, D. R. (1996). Towards a Minimalist Theory of Syntactic Structure. In H. Bunt	Constituency	18th International Conference on Computational Linguistics (COLING).	Constituency
the Association for Computational Linguistics (EACL-03). Budapest, Hungary,	Computing variation in gramm	& A. van Horck (eds.), Discontinuous Constituency, Berlin and New York, NY:	Computing variables in gramm	Saarbrücken, Germany: ICCL.	Computing validities in gramm
pp. 107-114. http://ling.osu.edu/~dm/papers/dickinson-meurers-03.html.	Amate	Eskin E (2000) Automatic Corrus Correction with Anomaly Detection. In	Amate	MS-CIS-05-09. Department of Computer Science and Information Science.	Amate
Dickinson, M. & W. D. Meurers (2003b). Detecting Inconsistencies in Treebanks. In	Garantinuity Computing car. rugama	Proceedings of the First Conference of the North American Chapter of the	Gaussinuty Computing our in-plants	University of Pennsylvania, Philadelphia.	Descripting on regions
Proceedings of the Second Workshop on Ireebanks and Linguistic Theories (TLT.03) Vävi5 Sweden on 45–58	Penalty for TATAN Increasing small	Association for Computational Linguistics (NAACL-00). Seattle, Washington.	Paraulto for TADER Increasing second	Huck, G. (1985). Exclusivity and discontinuity in phrase structure grammar. In	Parada to TADER Increasing serial
http://ling.osu.edu/-dm/papers/dickinson-meurers-tit03.html.	Parada	http://www.cs.coumbia.edu/~eeskin/papers/treebank-anomary-naaciuo.ps.	Family	West coast conterence on Formal Linguistics (WCCFL). Stanford University, CSU Publications, vol. 4, no. 92–98.	Family
Dickinson, M. & W. D. Meurers (2004). Error detection with discontinuous	Dependency Version exertion	(2004). Towards a Dependency-Based Gold Standard for German Parsers.	Dependency Visiolity depting	Huck, G. & A. Ojeda (eds.) (1987). Discontinuous Constituency. No. 20 in Syntax	Dependency Visitin electric
constituents. In P. Rodrigues, D. Cavar & J. Henring (eds.), Proceedings of the	Nature of Approximation	The TIGER Dependency Bank. In S. Hansen-Schirra, S. Oepen & H. Uszkoreit	Nature of dependencies	and Semantics. New York, et al.: Academic Press.	Natura di Aspendencian Informaziona
Pirst Mowest Computational Englistics Collegium: Biomington, Indiana.	investing small	(eds.), 5th International Workshop on Linguistically Interpreted Corpora (LINC-04) at COLING. Geneva, Switzerland: COLING, no. 31,-38, LIBI	increasing small	Jonansson, S. (1986). The Tagged LOB Corpus: Users' Manual. Norwegian Computing Contro for the Humaniton. Barran	investig seal
Language Corpora. In The Special Session on treebanks for spoken language	Summary	http://aclweb.org/anthology/W04-1905.	Summary	Kathol, A. (1995). Linearization-Based German Syntax. Ph.D. thesis. Ohio State	Summary
and discourse at NODALIDA-05. Joensuu, Finland. URL	UNIVERSITAT 9	Hajič, J. (2004). Disambiguation of Rich Inflection (Computational Morphology of	UNIVERSITAT P	University, Columbus, OH. Revised version published by Oxford University	UNIVERSITAT P
http://ing.osu.edu/~dckinso/papers/dickinson-meurers-nodaiidau5.ntml	PI/PI	Czech). Karolinum, Charles University Press, Prague, Czech Hepublic.	PI/PI	Press, 2000.	PI/PI
Kinn T.H. B. Cmurch S. Biezler M. Dalrymole & B.M. Kanlan (2003). The B&BC		McCawley J D (1982) Parentheticals and discontinuous constituent structure		P Mauther B Moullok & K Tauder (eris.) Text Shearth and Dialingue 4th	
700 Dependency Bank. In Proceedings of the 4th International Workshop on	Corpus Annotation	Linguistic Inquiry 13(1), 91–106.	Corpus Annotation	International Conference, TSD 2001, Zelezna Ruda, Czech Republic,	Corpus Annotation
Linguistically Interpreted Corpora, held at the 10th Conference of the European	Extractileuron University of Tablepon	Meurers, W. D. (2005). On the use of electronic corpora for theoretical linguistics.	Extractileurors University of Tabinger	September 11-13, 2001, Proceedings. Springer, vol. 2166 of Lecture Notes in	Extract Means University of Tablysen
URL http://www2.parc.com/is/inrouns/oft/fshank/	Introduction	Case studies from the syntax of German. Lingua 115(11), 1619-1639.	Introduction	Computer Science, pp. 39–46. Padro I. & I. Marnuez (1998). On the Evaluation and Comparison of Tanners: the	Introduction
Kroch, A. S. & A. K. Joshi (1987). Analyzing Extraposition in a Tree Adjoining	Ellarity of American Errors	mporngosuleau ampapersmeare/S-CC.mml.	Ellarity of Annulation Errors	Effect of Noise in Testing Corpora. In COLING-ACL. pp. 997-1002. URL	Ellarite et Armatudian Errora
Grammar. In Huck & Ojeda (1987).	Part of Speech	Initize, I. H. (2002). Category induction inom distributional cues in an artificial language. Memory & Cognition 30, 678–686.	Part of Speech	citeseer.ist.psu.edu/padro98evaluation.html.	Part of Speech
Kromann, M. T., L. Mikkelsen & S. K. Lynge (2004). Danish Dependency Treebank: Amentation Guide. http://www.id.ebs.dk/.cr/tk/treebank/suideT.html	Variation datastion	Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child	Variation detaction	Penn, G. (1999). Linearization and WH-extraction in HPSG: Evidence from South Constitut In P. D. Postley 8 A. Propeiddowski (eds.). Stavio in HPSG	Variation detection
Květěn, P. & K. Oliva (2002). Achieving an Almost Correct PoS-Tapped Corpus. In	Independent acidence from	directed speech. Cognition 90, 91–117.	Independent autiliance from	Stanford, CA: CSLI Publications, pp. 149–182.	Integrated editors from
P. Sojka, I. Kopeček & K. Pala (eds.), Text, Speech and Dialogue 5th	Annalis for the MSJ	Morrill, G. V. (1995). Discontinuity in categorial grammar. Linguistics and	Results for the MEL	Plätek, M., T. Holan, V. Kuboň & K. Oliva (2001). Word-Order Relaxations and	Analy in the MU
International Conference, TSD 2002, Brio, Czech Republic, September 9-12, 2002, Heidelbern: Springer on 2448 in Letture Notes in Artificial Intelligence.	Constituency	missiphy io, 1/5-219.	Constituency	Hestrictions within a Dependency Grammar. In G. Satta (ed.), Proceedings of the Seventh International Workshop on Parsing Technologies (IMPT). Builder	Constituency
(LNAI), pp. 19-26.	Variation datasetien	revising POS tags at the same time. In Proceedings of COLING.	Variation detailion	Tsinghua University Press, pp. 237-240.	Variation detaction
Leech, G. (1997). A Brief Users' Guide to the Grammatical Tagging of the British	William study	http://ling.osu.edu/~dm/02/spring/795K/mueller-ule.ps.	WELL new windy	Rambow, O. & A. Joshi (1994). A Formal Look at Dependency Grammars and	Will new study
National Corpus. UCREL, Lancaster University, Lancaster. http://www.bou.org.oo.uk/IBNC.industriencentrep.html	Gaussinaly	Müller, S. (1999). Deutsche Syntax deklarativ. Head-Driven Phrase Structure	Gamminuty	Phrase-Structure Grammars, with Special Consideration of Word-Order	Gaussinaly
Lezius, W. H. Biesinger & C. Gerstenberger (2002). TIGERRegistry Manual. IMS.	Computing our ingrams Paralis for TAXER	Grammar für das Deutsche. No. 394 in Linguistische Arbeiten. Tübingen: Max	Computing van rugsams Results for TADER	Environmena. In L. warmer (ed.), convert issues in Meaning-rext-Theory, London: Pinter, http://arxiv.org/abs/cmp-lo/9410007.	Computing van rugsams Paraulte for TADER
University of Stuttgart.	Property and	Nemeral I Marca A 1114 (2000) Televalue 27: A Consider Technologie	Pressing small	Reape, M. (1993). A Formal Theory of Word Order: A Case Study in West	Pressing serail
MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk. Mahwah,	Dependency	Phrase Structure and Dependency Annotation. In Proceedings of the fifth	Dependency	Germanic. Ph.D. thesis, University of Edinburgh, Edinburgh.	Dependency
NJ: Lawrence Eribaum Associates, third edition ed. Mannue M. B. Santariai & M. A. Manniekiawicz (1992). Building a large associated	Variation detection	international conference on Language Resources and Evaluation (LREC2006).	Variation detailion	Highter, H. & M. Saller (2001). On the Left Periphery of German Finite Sentences.	Variation datasetion
corpus of English: The Penn Treebank. Computational Linguistics 19(2),	Indust annalation	Genoa, Italy.	below analysis	Syntax, Stanford, CA: CSLI Publications, Studies in Constraint-Based	Indust annual days
313-330. ftp://ftp.cis.upenn.edu/pub/treebank/doc/ct93.ps.gz.	G many	Ojeda, A. (1987). Discontinuity, multidominances and unbounded dependency in Generalized Rhmon Structure Generatory. In Hurl: 8 Olecto (1997)	General	Lexicalism, pp. 257-300.	G many
Marcus, M., B. Santorini, M. A. Marcinkewicz & A. Taylor (1999). Treebank-3 Comut. Liepuistic Data Controllium. Bhiladeibhia		Olive K (2001) The Reschilling of Automatic Detection Computing of Errors in		Sampson, G. & A. Babarczy (2003). Limits to annotation precision. In Proceedings of the 4th International Mexichae on Lineuistically Internated Corport	
http://www.idc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC99T42.	TUNNGIN	Tagged Corpora: A Pilot Study on a German Corpus. In V. Matoušek,	TUNNGIN	(LINC-03). pp. 61–68. http://www.grsampson.net/Alta.html.	TUNNGIN
	91/91		91/91		91/91

Bentersin, R. (1980). Periol C System Diagray Guberines for the New Network Network of State (1980). Periol C System (1986). Science State Target acknowly The Act S. Tarker, T. Shari, N. M. Sci., Line: Bangara et al. State (1987). Science State (1987). Science State (1987). Science State (1987). Science State (1987). Science State (1988). Science State (1987). Science State (1987). Science State (1987). Science State (1987). Science State (1987). Science State State (1987). Science State (1987). Science State (1987). Science State State (1987). Science State (1987). Science State (1987). Science State (1997). Science State (1997). Science State (199	<section-header></section-header>	Muchann, A. T. Schwen (1996). Specifying a native granitational processing of the specific system of the specific system of the PLP (ACC Code). The specific system of the specific system of the process structures in the location specific distributions to a prover a displace of the specific structures. The specific system of the specific structures are applied to the specific structures of the specific system of the specific structures. The specific structures are applied by the specific system of the specific structures are applied to the specific structures. The specific system of the specific system of the specific structures of the specific system of the specific system of the specific structures. The specific system of the specific system of the specific structures of the specific system of the specific system of the specific structures of the specific system of the specific s	Control of the second s	
---	-----------------------------------	---	--	--