EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Universität Stuttgart

Philosophische Fakultät
Seminar für Sprachwissenschaft

Sonderforschungsbereich 732
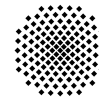Institut für Maschinelle Sprachverarbeitung

Focus, Givenness and Information Status
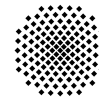## Annotating Corpora with Information Structure
ESSLLI 2014

Kordula De Kuthy and Arndt Riester

August 19, 2014
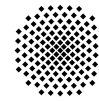
Questions to be addressed

- ► What is focus?

- ► What constitutes a focus theory?

- ► How can we describe the influence of focus on the meaning of sentences, and on their appropriateness in discourses?

- ► How can we detect focus in corpus data?

What is focus?

Answers provided in the literature:

- ▸ Focus is the **answer to a question** (to an explicit but also to an implicit one).
- ▸ Focus is the **informative** part of an utterance.
- ▸ Focus is the part of an utterance that signals **alternatives**.
- ▸ Focus indicates **new, or important** / **contrastive** information.
- ▸ Focus is **asserted** / **at issue**.
- ▸ Focus is often signalled by **prosodic or syntactic prominence**. *(language-dependent)*
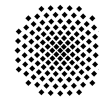
Two important theoretical contributions in the past

(1)   a.   Who is laughing?

        b.   JOHN*focus* [is laughing]*given/background*.

▶ Of the two most influential focus frameworks in the past 30 years, one concentrates on the focus part, the other on the given part.

▶ Mats Rooth's **Alternative Semantics (Rooth 1985, 1992, 1996, 2010)** is based on idea that focus triggers (contrastive) alternatives.

▶ Roger Schwarzschild **(Schwarzschild 1999)** develops a technical givenness notion.

▶ Contemporary theories of information structure, such as Büring (2008); Beaver & Clark (2008); Wagner (2012) and others, mainly build on, and combine, ideas from Rooth and Schwarzschild.
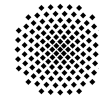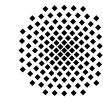
Schwarzschild (1999): GIVENness, AvoidF and other constraints on the placement of accent

Prerequisite: focus projection (discussed yesterday)

- *Focus* is not equivalent with the word that carries the *pitch accent*.

- In West-Germanic languages (English, Dutch, German), focus originates in an accented word, and then *projects* onto larger phrases (Gussenhoven 1983, 1992, 1999; Rochemont 1986; Selkirk 1984, 1995; Winkler 1997).

(2)   [MAry$_f$ [bought$_f$ [a$_f$ [book$_f$ [about$_f$ BATS$_f$]$_f$]$_f$]$_f$]$_f$]$_F$

Limits of focus projection – I

- Focus projection is not mandatory.
- Normally, focus projects in order to achieve discourse congruence.
- But focus projection cannot solve all problems.
- Focus does not project from the head of a phrase to an argument.

(3) $[\text{MAry}_f \,[\text{bought}_f \,[\text{a}_f \,[\text{BOOK}_f \,[\text{about bats}]]_f]_f]_f]_F$.

- In sentence (3) *about bats* cannot get an F-mark.
- This can become problematic.
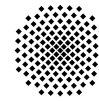
Givenness principle (Schwarzschild 1999)

- **Constituents which are not F-marked, must be given**.
- Reversal: constituents which are not given must be F-marked.
- They must either be accented or "borrow" an F-mark by means of focus projection.

(14) $[\text{MAry}_f\ [\text{bought}_f\ [\text{a}_f\ [\text{BOOK}_f\ [\text{about bats}]]_f]_f]_f]_F$.

- I.e. (3) can only be used in a context which already talks about *bats*. Otherwise it will be infelicitous.
- Caution: the givenness principle does not imply that given constituents must not be F-marked.

(15)   a.   Do you prefer vanilla or walnut?
       b.   I prefer WALnut_F.   (given but F-marked)
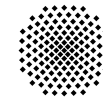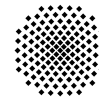
Limits of focus projection – II

▶ Modifiers are typically adjuncts (optional information), not arguments. F-marks on modifiers do not project to the head noun.

(16) {What did John drive?}

  a. #He drove a $BLUE_F$ convertible.
  Givenness principle violated, because the F cannot project onto "convertible", which is new.

  b. ✓He drove [the$_f$ [convertible$_f$ [of$_f$ [his$_f$ MUM$_f$]$_f$]$_f$]$_f$]$_F$.
 *("convertible" receives f via horizontal projection)*

Schwarzschild's goal

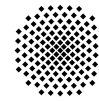Provide a unified theory that accounts for the accent patterns in the following cases:

(17) a. Why don't you have some French toast?
   b. I've forgotten how to MAKE French toast. *(newness)*

(18) a. John's mother voted for Bill.
   b. No, she voted for JOHN. *(contrast / correction)*

(19) a. Who did John's mother vote for?
   b. She voted for JOHN. *(question-answer)*

▶ Halliday (1967) redefines *newness* to capture all these cases. This is not intuitive.

▶ Schwarzschild: a unified theory should not make reference to new information at all.

▶ What we need to do is redefine *GIVENness*.
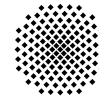
Phonological observations on English

- Prominence indicates novelty.   **Wrong!**
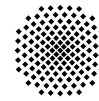
- Lack of prominence indicates givenness.   **Correct!**

AVOIDF

- GIVENNESS: Constituents which are not F-marked must be given.
- In order to avoid a violation of GIVENNESS, we might simply F-mark everything, e.g. place a pitch accent everywhere.
- But this is not what is happening.
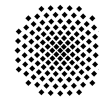- There must be an additional constraint which tells us to use accents sparingly: AVOIDF

AVOIDF:

F-mark as little as possible, without violating GIVENNESS.

Question-answer congruence (Halliday 1967)

- An appropriate answer to a wh-question must have F-marking on the constituent corresponding to the wh-phrase.

(20)   a.   What did Mary do?
       b.   She [praised$_f$ [her$_F$ BROTHER$_f$]$_f$]$_F$.

(21)   a.   What did John's mother do?
       b.   She [[PRAISED$_f$ him]$_F$.

(22)   a.   Who did John's mother praise?
       b.   She praised HIM$_F$.

Consequences of Schwarzschild's approach

Recall:

(23)   What did John drive?
        #He drove a BLUE$_F$ convertible.

- ▸ GIVENNESS violated.

(24)   John drove Mary's red <u>convertible</u>. What did he drive
        before that?
            ✓He drove a BLUE$_F$ convertible.

- ▸ No GIVENNESS violation
- ▸ Question-answer congruence is lost on Schwarzschild's
  account.

What does *given* mean after all?
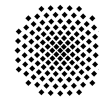
(25)   John drove Mary's red convertible. What did he drive before that?

     a.   He drove [a BLUE$_F$ convertible].

- The phrase [a BLUE$_F$ convertible] is not F-marked itself. Hence it must be *given*. But is it?
- The indefinite phrase introduces a new referent into the discourse (Heim 1982; Kamp & Reyle 1993).
- Also intuitively, since it contains new material, the phrase is not entirely given.
- The same goes for the entire sentence.
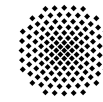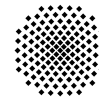
GIVEN (Schwarzschild 1999)

An utterance U counts as GIVEN iff it has a a salient antecedent A and

a. *if U is type e, then A and U co-refer;*
b. *otherwise: modulo $\exists$-type shifting, A entails the $\exists$-F-closure of U.*

What does that mean?

- ▶ If an expression is of type *e* (e.g. *he*), it must be co-referential, e.g. with the earlier mentioned *Paul*.

- ▶ If an expression is of type $\langle \alpha, \beta \rangle$ (e.g. the verb *moves*), it must be entailed by some other expression in the discourse context (e.g. *walks*).

- ▶ How can we say that an arbitrary expression entails another one?

Existential Typeshift

$$\llbracket \textit{walks} \rrbracket \ = \ \lambda x.\textit{walk}(x)_{\langle e,t \rangle} \qquad \llbracket \textit{moves} \rrbracket \ = \ \lambda x.\textit{move}(x)_{\langle e,t \rangle}$$

► Type shift to proposition level: replace lambdas by existential quantifiers.

$$\exists x.\textit{walk}(x)_t \qquad \exists x.\textit{move}(x)_t$$

► Check whether the typeshifted antecedent entails the typeshifted "anaphor".
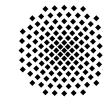
► If such an entailment relation can be established then *moves* is GIVEN.

Existential Typeshift

$$\llbracket walks \rrbracket = \lambda x.walk(x)_{\langle e,t \rangle} \quad \llbracket moves \rrbracket = \lambda x.move(x)_{\langle e,t \rangle}$$

▶ Type shift to proposition level: replace lambdas by existential quantifiers.

$$\exists x.walk(x)_t \models \exists x.move(x)_t \ \checkmark$$

▶ Check whether the typeshifted antecedent entails the typeshifted "anaphor".

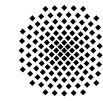▶ If such an entailment relation can be established then *moves* is GIVEN.

Existential F-Closure

- F-marks function as "wildcards" in the entailment process.
- The phrase *a BLUE$_F$ convertible* should count as GIVEN if it occurred after *a red convertible*.
- **∃-F-closure:** replace F-marked part of a constituent by an existentially bound variable *(an X-colored convertible)*.
- Perform existential typeshift (here: from quantifier to proposition).

$$\text{a red convertible} \models \text{an X-colored convertible}$$

$$\exists P \exists x[conv(x) \wedge red(x) \wedge P(x)] \models \exists P \exists X \exists x[conv(x) \wedge X(x) \wedge P(x)]$$
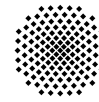
Predicting focus and accent in an OT-model

► Schwarzschild's model only allows for indirect predictions of focus and accent placement.

► A set of candidates with different F-distributions is generated.

► Each candidate is tested for its compliance with GIVENNESS.

► AVOIDF: If several candidates pass GIVENNESS, the one with the least F-marks is chosen.
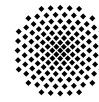
A very simple example

(26)   Whom did $John_1$'s $mother_2$ praise?
            She_2 praised him_1.

Candidates:

i. She praised HIM_F.
ii. She $[praised_f \ HIM_f]_F$.
iii. She $PRAISED_F$ him.
iv. She $[PRAISED_f \ him]_F$.
v. SHE_F praised him.
vi. SHE_F praised HIM_F.
vii. . . .

Checking for GIVENness

(26)   Whom did $John_1$'s $mother_2$ praise?

   i. She_2 praised $[HIM_1]_F$.

$[\![John's\ mother_2]\!]$ $\leftrightarrow$ $[\![she_2]\!]$ ✓
$[\![did \ldots praise]\!]$ $\models$ $[\![praised]\!]$ ✓
$[\![John_1]\!]$ $\leftrightarrow$ $[\![him_1]\!]$ ✓

▶ $[\![praised\ him_1]\!]$ is not GIVEN but F-closure saves the day.

$[\![did \ldots praise\ whom]\!]$ $\models$ $[\![praised\ HIM_F]\!]$ ✓

▶ Hence, candidate (i) is good (and minimally F-marked).

Other candidates?

- ▶ iii. She PRAISED$_F$ him.
- ▶ (iii) is also minimally F-marked, but. . .
- ▶ ⟦*PRAISED$_F$ him$_1$*⟧ is not GIVEN.
- ▶ Even after applying F-closure (replacing the verb by "did something") the context does not entail that "somebody did something to John".
- ▶ Hence, (iii) is not a good candidate.
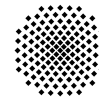
Convertible example

(27)   a.   John drove Mary's red convertible. What did he drive before that?

          b.   He drove her $BLUE_F$ convertible.

By the same reasoning, we can show that all constituents of (27b) are GIVEN.
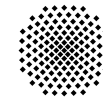
Upshot

- ▶ On Schwarzschild's account, question-answer congruence is lost.
- ▶ Computations for natural data can become extremely complex.
- ▶ Summary of the account: *If F-marks are distributed in an appropriate manner, then every constituent is – technically – GIVEN.*
- ▶ Very unintuitive givenness notion
- ▶ The account may get almost all examples right, but it is not really useful for annotation purposes.
- ▶ In the following, we will develop another approach to givenness that follows a different tradition but integrates some of Schwarzschild's ideas.
- ▶ First, back to the basics of *givenness*...
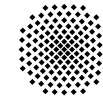
Information status

Information status of referring expressions

- Information status – originally – describes the **degree of givenness** / **salience** / **cognitive activation** / **accessibility** of referring expressions.
- Goal: distinguish classes of expressions in text or spoken discourse in a way that is as **fine-grained** as possible and still **reproducible** by non-experts with **high reliability**
- Prince (1981) distinguishes **textually** / **situationally evoked**, **inferrable** and **new** expressions.
- Prince (1992) introduces two dimensions (discourse status vs. hearer status):

	hearer-old	**hearer-new**
discourse-old	given, old, active	–
discourse-new	unused, familiar, known	brand-new

Information status

- ► Cognitive activation: Chafe (1994) distinguishes highly salient **(active / given / consciously available)**, less salient **(semi-active / accessible / unconscious)**, and non-salient **(inactive / new)**

- ► Other notable classifications (each one with their own use of terminology) are Gundel et al. (1993) **(givenness hierarchy)**, Lambrecht (1994); Poesio & Vieira (1998); Eckert & Strube (2000); Ariel (2001) **(accessibility theory)**, Nissim et al. (2004); Götze et al. (2007); Riester et al. (2010)

- ► An overview and partial comparison provided in Baumann & Riester (2012)

Our system

- *RefLex* scheme for the annotation of information structure (Baumann & Riester 2012; Riester & Baumann ms.)
- Goal: combine the approach by Schwarzschild (1999) with earlier accounts of information status
- Address a number of problems of earlier accounts
- Enable annotations on natural data that are both reliable and fine-grained
- Two levels:
  1. Referential Information Status *(referring expressions)*
  2. Lexical Information Status *(non-referring expresions)*

Analysis of referring expressions: historic perspective

▶ The analysis of referring expressions has a long history in linguistics and philosophy.

▶ Two early claims by Frege (1891, 1892) on *definite descriptions* and *proper nouns / names*:

1. The felicitous use of a definite *presupposes* the existence of an entity to which the definite can refer.
2. This presupposed entity is *unique*, i.e. there is exactly one entity that satisfies the description.

▶ This is indeed the case for certain definites: *the sun, the present Pope, Gottlob Frege, the President of the United States, the positive square root of 4, . . .*
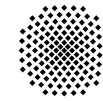
Analysis of referring expressions (cont.)

- ▶ Russell (1905): Quantificational analysis of definite descriptions
- ▶ Strawson (1950): Criticism of Russell, restoring a variant of Frege's referential account
- ▶ There is something wrong about the uniqueness assumption in definites like *the table, the cup, the man, the ant, the molecule, . . .*
- ▶ Uniqueness must be relativized to different *types of contexts*.
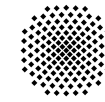
Types of context

Definite descriptions can be unique with respect to different context types:

Context	Label (RefLex)	Phenomenon
previous discourse context	*r-given*	coreference anaphora
communicative situation	*r-given-sit* *r-environment*	symbolic deixis gestural deixis
frame / **scenario**	*r-bridging*	bridging / associative anaphora
following discourse context	*r-cataphor*	cataphora
global context	*r-unused*	global uniqueness

An (indefinite) expression, which refers non-uniquely, receives the label *r-new*.
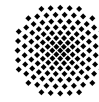
Annotation conventions for referential expressions

▸ Annotate all referring expressions (i.e what is called **DP** in generative linguistics, or **NP** in e.g. computational linguistics):

(28)   *a cat, she, his, the table, this ugly lamp, John, Chancellor Merkel, Eddie's, Tübingen, someone, freedom, squirrels, the guy who is sleeping* etc.

▸ Quantified DPs:

(29)   *every participant, many suitcases, a lot of work, few factories* etc.

▸ In case a preposition is present, it is included in the markable:

(30)   *(asked) for the bill, (went) to Tunisia, because of the new law, with several friends* etc.
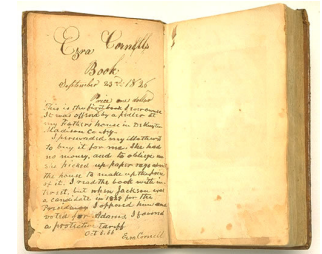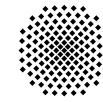
Annotation conventions (cont.)

- ▶ Appositions are included:

(31)   John Smith, the ambassador; a drink, which turned out to be mango lassi; Harry, who hasn't been seen for two weeks

- ▶ Focus-sensitive particles are *not* included in the markable:

(32)   *(only) a snack; (even) Helga; the assignment(, too)*

r-given **(old, active, textually evoked)**

Coreference, uniqueness in previous discourse

(33)   I met <u>a friend</u> yesterday.

    a.   [He] told me a story.   *(pronoun)*

    b.   [The friend] came from Hamburg.   *(same noun)*

    c.   [The old chap] was very tired.   *(different expression)*

    d.   I hadn't seen [Albert] for months.   *(name)*

(34)   The West is suspecting <u>Iran</u> of building nuclear arms. But negotiations with [Teheran] continue.

                                         (metonymy / synecdoche)

(35)   [Paul [sings under the shower]$_k$]$_i$

    a.   Mary finds [that]$_i$ weird.

    b.   John does [it]$_k$, too.   *(abstract anaphora)*
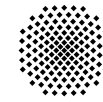
r-given-sit (situationally evoked)

Symbolic deixis, uniqueness relative to communicative situation

(36)   [I] want [us] to return the car.

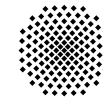(37)   [Last week], she told [me] the opposite.

(38)   Come [here]!

We do not annotate temporal quantifiers like *always, often, every Wednesday*

Snowden interview: deixis and anaphora

218 [01:21.5]	219 [01:	220 [01:	221 [01:	222 [01:	223 [01:22.	224 [01:2	225 [01:	226 [01:	227 [01:24.4]	228 [01:	229 [01:	230 [01:24.8]	2
You	can	wire	tap	a	Federal	Judge	and	if	you	do	it	carefully	r
PP	MD	VB	VB	DT	NP	NP	CC	IN	PP	VBP	PPRB	RB	F
r-given-sit +generic									r-given-sit +generic		r-given		

- ▶ Abstract anaphor: antecedent of *it* highlighted
- ▶ Additional feature +*generic* marks that the expression does not refer to a specific individual but has a class reading.

r-environment

- Gestural deixis, uniqueness of visual object ensured by demonstration
- Occurs only in face-to-face communication
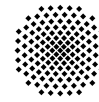
(39)   You should take [this way].

(40)   [He] kicked me.

r-bridging (associative, mediated)

- ▸ (Clark 1977; Asher & Lascarides 1998; Poesio & Vieira 1998; Löbner 1998)
- ▸ Discourse-new but dependent on previous context
- ▸ Uniqueness within scenario / frame
- ▸ An expression with an implicit argument
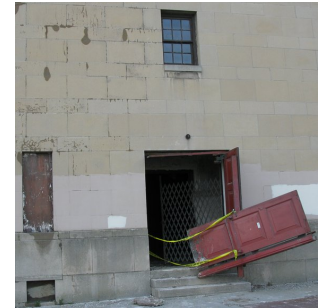
(41)   When they tried to enter the house, the door fell off.

(42)   The city is planning a new townhall and [the construction] will start early next year.

(43)   Our correspondent in Egypt is reporting that [the opposition] is holding a rally [against the constitutional referendum].

Note that bridging is not *defined* in terms of a part-whole relation!
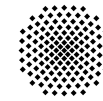
Snowden interview: bridging anaphora

182 [01:0	183 [01:10.4]	184 [01:	185 [01:	186 [01:	187 [01:11.3]	188 [01::	189 [01:12.3]	190 [01:	191 [01:	192 [01:	193 [01:12.9]
When	you	are	on	the	inside,	when	you	go	into	work	everyday,
WRB	PP	VBP	IN	DT	NN	WRB	PP	VBP	IN	NN	JJ
	r-given-sit +generic		r-bridging				r-given-sit +generic		r-bridging		
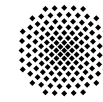
▶ Bridging anaphors have an antecedent (sometimes silent) that is understood as an implicit argument.

(44)   on the inside (of the NSA)

(45)   into (your) work

▶ Assign a label to an entire phrase:

r-unused-known

▸ Unique expression in the global context (first mention)

▸ Likely to be known by the intended audience
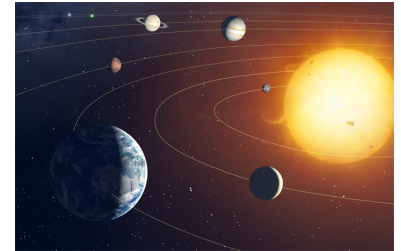
(46)   [The Pope] stood [on St. Peter's Square].

(47)   [Space probe Voyager 1] passed [planet Jupiter] [in 1979].

(48)   [Igor Stravinsky] died [in New York] and was buried [in Venice].

▸ Note that the question whether an entity is *known* by the audience is not a linguistic question but varies over time and for different addressees.

▸ There are also globally unique entities which are *unknown*.
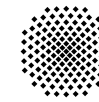
Snowden interview: known expression

22 [00:10.0]	23 [00:1(	24 [00:10.6]	25 [00:1	26 [00:1	27 [00:11.5]	28 [00:1	29 [00:1:	30 [00:1:	31 [00:1:	32 [00:12.	33 [00:1:	34 [00:13.4*]
You	were	working	till	last	summer	for	the	NSA	and	during	this	time
PP	VBD	VBG	IN	JJ	NN	IN	DT	NP	CC	IN	DT	NN
r-given-sit				r-given-sit		r-unused-known				r-given		

Complex phrases

- Often, several referring expressions are nested inside each other.
- Each of them receives its own label.

(49)   [All activities [on the international airport [in the vicinity]]] came to a halt.

- The same goes for possessive pronouns.

(50)   He welcomed them [to his house].
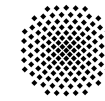
Annotation on available syntactic structures

DIRNDL corpus (Eckart et al. 2012; Björkelund et al. 2014), SALTO tool (Burchardt et al. 2006)

© 2014 Universität Tübingen, Universität Stuttgart
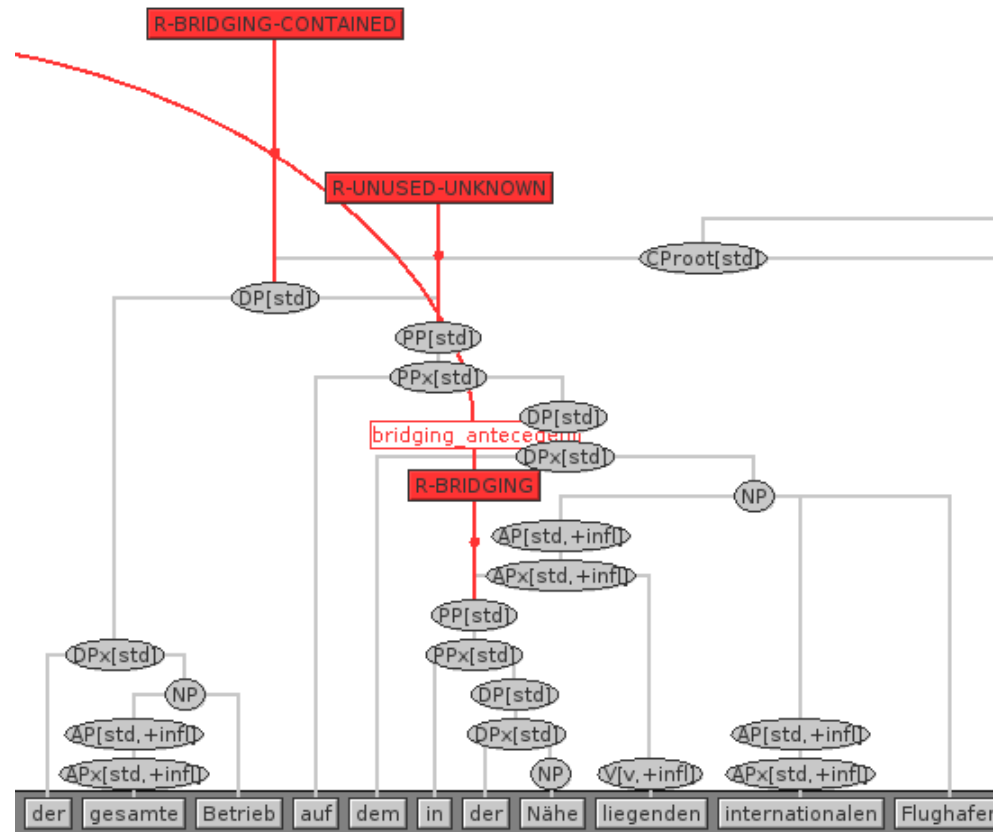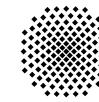
Complex expressions in EXMARaLDA

`snowden-interview-answer8-snowden-anno.exb`

13 [00:0:	14 [00:0:	15 [00:0:	16 [00:04	17 [00:04_5]	18 [00:05,0]	19 [00:0.	20 [00:05,4]	21 [00:0.	22 [00:0.	23 [00:0:	24 [00:0:	25 [00:06.7*]	26 [00
and	point	to	policy	agreements	between	the	members	of	the	Five	Eyes	saying	that
CC	NN	TP	NN	NNS	IN	DT	NNS	IN	DT	CD	NNS	VBG	IN
		r-unused-unknown											
					r-bridging-contained								
								r-given					

▶ Create a label tier for each level of embedding.
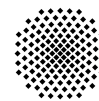
r-unused-unknown

- ► Unique expression in the global context (first mention)
- ► Not likely to be known by the audience
- ► Sufficient descriptive material to ensure identifiability and to introduce a unique new entity to the common ground

(51)   [The woman Max went out with last night] is an astrophysicist.

(52)   [The swimming pool of the new town hall] created discontent among the voters.

(53)   I just saw [the creepy reptile of my office colleague].

(54)   [Carl, my neighbour,] never gets up before 10 o'clock.

r-bridging-contained containing inferrable

- ▸ Special type of bridging anaphor: the anchor / antecedent is a syntactic argument of the head noun, i.e. it is embedded in the phrase.

(55)   [The opening day of the G20 summit] was a desaster.

(56)   We were surprised to even see [the President of Malta].

(57)   [The construction of the new townhall] will start early next year.

Distinguishing r-unused-unknown and r-bridging-contained

Permutation test: if there is an embedded argument, front it. If the resulting discourse is felicitous, assign the label *r-bridging-contained*. Otherwise, assign *r-unused-unknown*.

(58)   a.   Markable: [the President of Malta]

b.   Permutation: ✓ I was in Malta and met the President.
⇒ *r-bridging-contained*

(59)   a.   Markable: [the creepy reptile of my office colleague]

b.   Permutation: ??When my office colleague left the room, [the creepy reptile] attacked.
⇒ *r-unused-unknown*

r-new

- Non-unique, discourse-new expression

(60)   [One stormtrooper] threatened me.

(61)   [A military spokesman] confirmed [explosions] and the death [of at least two soldiers].

(62)   I'm married [to a computer scientist].

References

Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schliperoord & W. Spooren (eds.), *Text representation: Linguistic and psycholinguistic aspects*, Amsterdam: Benjamins, pp. 29–87.

Asher, N. & A. Lascarides (1998). Bridging. *Journal of Semantics* 15, 83–113.

Baumann, S. & A. Riester (2012). Referential and Lexical Givenness: Semantic, Prosodic and Cognitive Aspects. In G. Elordieta & P. Prieto (eds.), *Prosody and Meaning*, Berlin: Mouton de Gruyter, vol. 25 of *Interface Explorations*, pp. 119–162.

Beaver, D. & B. Clark (2008). *Sense and Sensitivity. How Focus Determines Meaning*. Chichester: Wiley & Sons.

Björkelund, A., K. Eckart, A. Riester, N. Schauffler & K. Schweitzer (2014). The Extended DIRNDL Corpus as a Resource for Coreference and Bridging Resolution. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC)*. Reykjavik.

Burchardt, A., K. Erk, A. Frank, A. Kowalski & S. Padó (2006). SALTO: A Versatile Multi-Level Annotation Tool. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC)*. Genoa, Italy.

Büring, D. (2008). What's New (and What's Given) in the Theory of Focus? In *Proceedings of the 34th Annual Meeting of the Berkeley Linguistics Society*. pp. 403–424.

Chafe, W. L. (1994). *Discourse, Consciousness, and Time*. University of Chicago Press.

Clark, H. H. (1977). Bridging. In P. Johnson-Laird & P. Wason (eds.), *Thinking: Readings in Cognitive Science*, Cambridge University Press, pp. 169–174.

Eckart, K., A. Riester & K. Schweitzer (2012). A Discourse Information Radio News Database for Linguistic Analysis. In C. Chiarcos, S. Nordhoff & S. Hellmann (eds.), *Linked Data in Linguistics. Representing and Connecting Language Data and Language Metadata*, Heidelberg: Springer, pp. 65–76.

Eckert, M. & M. Strube (2000). Dialogue Acts, Synchronising Units and Anaphora Resolution. *Journal of Semantics* 17(1), 51–89.

Frege, G. (1891). *Function und Begriff: Vortrag gehalten in der Sitzung vom 9. Januar, 1891 der Jenaischen Gesellschaft für Medicin und Naturwissenschaft*. Jena: Hermann Pohle.
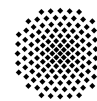
Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100, 25–50.

Götze, M., T. Weskott et al. (2007). Information Structure. In S. Dipper, M. Götze & S. Skopeteas (eds.), *Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics and Information Structure*, Universitätsverlag Potsdam, vol. 7 of *Interdisciplinary Studies on Information Structure*. SFB 632.

Gundel, J. K., N. Hedberg & R. Zacharski (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language* 69(2), 274–307.

Gussenhoven, C. (1983). Focus, Mode and Nucleus. *Journal of Linguistics* 19, 377–417.

Gussenhoven, C. (1992). Sentence Accents and Argument Structure. In I. Roca (ed.), *Thematic Structure: Its Role in Grammar*, Berlin: Foris, pp. 79–106.

Gussenhoven, C. (1999). On the Limits of Focus Projection in English. In P. Bosch & R. van der Sandt (eds.), *Focus. Linguistic, Cognitive and Computational Perspectives*. Cambridge University Press, pp. 43–55.

Halliday, M. (1967). Notes on Transitivity and Theme in English. Part 2. *Journal of Linguistics* 3, 199–244.

Heim, I. (1982). The Semantics of Definite and Indefinite Noun Phrases. Ph.D. thesis, University of Massachusetts, Amherst.

Kamp, H. & U. Reyle (1993). *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer.

Lambrecht, K. (1994). *Information Structure and Sentence Form*. Cambridge University Press.

Löbner, S. (1998). Definite Associative Anaphora. In *Proceedings of the Discourse Anaphora and Resolution Colloquium (DAARC)*. Lancaster.

Nissim, M., S. Dingare, J. Carletta & M. Steedman (2004). An Annotation Scheme for Information Status in Dialogue. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC)*. Lisbon.

Poesio, M. & R. Vieira (1998). A Corpus-Based Investigation of Definite Description Use. *Computational Linguistics* 24(2), 183–216.

Prince, E. F. (1981). Toward a Taxonomy of Given-New Information. In P. Cole (ed.), *Radical Pragmatics*, New York: Academic Press, pp. 233–255.

Prince, E. F. (1992). The ZPG Letter: Subjects, Definiteness and Information Status. In W. C. Mann & S. A. Thompson (eds.), *Discourse Description: Diverse Linguistic Analyses of a Fund-Raising Text*, Amsterdam: Benjamins, pp. 295–325.

Riester, A. & S. Baumann (ms.). RefLex Scheme – Annotation Guidelines. Version of August 1, 2014.

Riester, A., D. Lorenz & N. Seemann (2010). A Recursive Annotation Scheme for Referential Information Status. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC)*. Valletta, Malta, pp. 717–722.

Rochemont, M. S. (1986). *Focus in Generative Grammar*. Amsterdam: Benjamins.

Rooth, M. (1985). Association with Focus. Ph.D. thesis, University of Massachusetts, Amherst.

Rooth, M. (1992). A Theory of Focus Interpretation. *Natural Language Semantics* 1(1), 75–116.

Rooth, M. (1996). Focus. In S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, Oxford: Blackwell, pp. 271–297.

Rooth, M. (2010). Second Occurrence Focus and Relativized Stress F. In C. Féry & M. Zimmermann (eds.), *Information Structure: Theoretical, Typological, and Experimental Approaches*, Oxford University Press.

Russell, B. (1905). On Denoting. *Mind* 14, 479–493.

Schwarzschild, R. (1999). GIVENness, AvoidF, and Other Constraints on the Placement of Accent. *Natural Language Semantics* 7(2), 141–177.

Selkirk, E. (1984). *Phonology and Syntax*. Cambridge: MIT Press.

Selkirk, E. (1995). Sentence Prosody. Intonation, Stress, and Phrasing. In J. Goldsmith (ed.), *The Handbook of Phonological Theory*, Blackwell, pp. 550–569.

Strawson, P. F. (1950). On Referring. *Mind* 59, 320–344.

Wagner, M. (2012). Focus and Givenness: a Unified Approach. In I. Kučerová & A. Neeleman (eds.), *Contrasts and positions in information structure*, Cambridge University Press, pp. 102–148.

Winkler, S. (1997). *Focus and Secondary Predication*. Berlin: Mouton de Gruyter.