On the Dimensions of Discourse Salience

Christian Chiarcos
chiarcos@uni-potsdam.de
Dimensions of Salience

• Background
  • Models of salience-based information packaging
    referential choice, grammatical roles and word order

• Corpus study 1
  One or multiple dimensions of salience ?

• Corpus study 2
  Forward-looking vs. backward-looking salience ?

• Discussion
Background: Linguistic Variability

- `[E]s darf nicht verkannt werden, dass man denselben Sinn, *denselben Gedanken auch verschieden ausdrücken* kann, wobei denn also die Verschiedenheit ... nur eine der ... Färbung des[selben] Sinnes ist und für die Logik nicht in Betracht kommt.‘
  
  (Frege 1892)

  `[W]e must not fail to recognize that the same sense, the same thought, may be variously expressed; thus the difference does ... concern ... only the ... colouring of the [same] thought, and is irrelevant for logic.’

  (Geach and Black 1980)

⇒ Linguistic variability cannot be (completely) accounted for on grounds of (Fregean) semantics

⇒ “Information Packaging”
Information Packaging

• `the kind of phenomena ... that ... have to do primarily with how the message is sent and secondarily with the message itself’

  (Chafe 1976)

• `the linguistic dimension that allows speakers to make structural choices in accordance with their assumptions about the hearer’s communicative state, and that allows hearers to decode the import of those structural choices appropriately.’

  (Vallduví 1994)
Information Packaging

- (a) the noun may be either given or new;
- (b) it may be a focus of contrast;
- (c) it may be definite or indefinite;
- (d) it may be the subject of the sentence;
- (e) it may be the topic of the sentence;
- (f) it may represent the individual whose point of view the speaker takes, or with whom the speaker empathizes (Chafe 1976)
Information Packaging

- (a) the noun may be either given or new;  
  (Sgall et al. 1986; Prince 1981)

- (b) it may be a focus of contrast;  
  (Langacker 1997)

- (c) it may be definite or indefinite;  
  (Lewis 1979)

- (d) it may be the subject of the sentence;  
  (Fillmore 1977)

- (e) it may be the topic of the sentence;  
  (Sgall et al. 1986; Grosz et al. 1995)

- (f) it may represent the individual whose point of view the speaker empathizes  
  (Chafe 1976)

Many aspects of information packaging have been explained on grounds of „salience“
Information Packaging

(a) the noun may be either given or new; (Sgall et al. 1986; Prince 1981)
(b) it may be a focus of contrast; (Langacker 1997)
(c) it may be definite or indefinite; (Lewis 1979)
(d) it may be the subject of the sentence; (Fillmore 1977)
(e) it may be the topic of the sentence; (Sgall et al. 1986; Grosz et al. 1995)

(f) it may represent the individual whose point of view the speaker empathizes (Chafe 1976)

Many aspects of information packaging have been explained on grounds of "salience"... but what exactly is it, and what effects does it have?
Effects of salience?

- Well, different people have different ideas

- **Personal pronouns** are more salient than **demonstratives**
  
  (Gundel et al. 1993)

- **Demonstratives** are more salient than **personal pronouns**
  
  (Sgall et al. 1986)

- salient (**given**) precedes non-salient (**new**)
  
  (Sgall et al. 1986)

- **new(sworthy)** precedes **given**
  
  (Mithun 1993)

- the grammatical subject designates **salient referents**
  
  (Fillmore 1977)

- the grammatical subject designates **non-salient referents** that are to be promoted in their saliency
  
  (Mulkern 2007)
What is salience?

- Well, different people have different ideas
  - salient = given?  
    (Sgall et al. 1986, Prince 1981)
  - salient = new(sworthy)?  
    relevant/important?  
    (Davis & Hirschberg 1988, Steedman 2000)
  - multiple dimensions of salience?  
    - backward-looking vs. forward-looking  
What is salience?

- Well, different people have different ideas.

> As we have just seen, a number of different factors have been claimed to contribute to salience. Researchers are also divided on the effects of salience to sentences. ... [S]alience is (...) characterized by a number of superficially dissimilar properties.

(Sridhar 1988)
What is salience?

Well, different people have different ideas.

As we have just seen, a number of different factors have been claimed to contribute to salience. Researchers are also divided on the effects of salience to sentences. (Sridhar 1988)

... but it is generally accepted that:

- salience has to do with attention and memory
- salience plays a crucial role in selection tasks
- this includes the information packaging of discourse referents
  - referential choice: pronominal > nominal
  - grammatical roles: subject > object > oblique
  - word order: salient precedes non-salient
What is salience?

Salience of discourse referents

- Monodimensional
  - Discourse referents are characterized by a single cognitive dimension of salience that governs referential choice, grammatical roles and word order preferences

- Multidimensional
  - At least two logically independent dimensions of salience are to be distinguished. Both interact in the derivation of packaging preferences for referential choice, grammatical roles and word order preferences
Two views on salience of discourse referents

**Multidimensional**

- Salience factors
  - Anaphoric salience factors
    - based on shared knowledge, e.g., about the preceding discourse
  - Other salience factors
    - sensitive to speaker-private intentions, e.g., with respect to the subsequent discourse

**Forward-looking**

- Grammatical roles
- Referential choice
- Word order

**Backward-looking**

- Information Packaging

**Monodimensional**

- Salience factors
  - Anaphoric salience factors
  - Other salience factors

- Attentional states accessibility in memory

- Information Packaging

- Grammatical roles
- Referential choice
- Word order


*(Sgall et al. 1986, Tomlin 1995, 1997)*
Salience in discourse

- **Salience of discourse referents**
  - is related to the focus of attention and accessibility in memory of hearer and/or speaker
  - is manifested by the choice of referring expressions, grammatical roles and word order
  - is the most important cognitive determinant of information packaging

- **Monodimensional vs. Multidimensional**
  - No agreement as to whether salience is a unified cognitive concept
Salience in discourse

- Two corpus studies
  - Test predictions of both models for the correlation between salience-marking grammatical devices
    - Pronominalization
    - Sentence-initial word order
    - Subject role
  - Test whether the dimensions of salience correlate with forward-looking and backward-looking salience factors
One or two dimensions of salience?

- **Background**
  - Salience influences information packaging
    - pronominalization, subject role, sentence-initial position

- **Corpus study 1**
  - One or two dimensions of salience?

- **Corpus study 2**
  - Forward-looking vs. Backward-looking salience?

- Discussion
Salience in discourse

- Corpus study
- German
  - Grammatical roles and word order less dependent on each other than in English
- TüBa-D/Z  
  - 2,213 newspaper articles
  - Syntax + coreference annotation
- Features
  - perspron (personal pronoun)
  - sbj (subject role)
  - vf (vorfeld, sentence-initial topological field)
Salience in discourse

- **Feature extraction**
  - SWI Prolog conversion of TüBa-D/Z
  - non-coordinated, non-embedded main clauses
    40,713 clauses
  - all nominal and prominal arguments and adjuncts
    79,222 (potential) referring expressions
  - packaging phenomena
    - perspron $\leftrightarrow$ pos="PPER"
    - subj $\leftrightarrow$ func=/on|onk/
    - vf $\leftrightarrow$ cat=","VF"
  - discourse features
    - given $\leftrightarrow$ link* to preceding discourse
    - important $\leftrightarrow$ link* to subsequent discourse

* "coreferential“, "anaphoric“, "bound“, "cataphoric“ or "instance“ relation
One or two dimensions?

- **Monodimensional prediction**
  - Salience understood as a latent variable
    - Can be extrapolated from information packaging
    - Extrapolation is imprecise
      - other (semantic, socio-cultural, etc.) factors have an influence on the realization of the referent
    - Reliability of the extrapolation increases, if multiple dimensions of information are taken into consideration that indicate the same salience status
One or two dimensions?

• Monodimensional prediction

• Salience-marking grammatical devices
  • Pronominalization (perspron)
  • Subject role (sbj)
  • Sentence-initial position (vf)

• Prediction 1
  • salience has an effect on information packaging
    • sbj => salient => perspron
    ⇒ sbj => perspron preference

\[ P(\text{perspron} | \text{sbj}) > P(\text{perspron}) \]
One or two dimensions?

- Monodimensional prediction

- Prediction 2

\[ P(X_{\text{sal}} \mid Y_{\text{sal}}, Z_{\text{sal}}) \geq P(X_{\text{sal}} \mid Y_{\text{sal}}) \]

- Salience extrapolation from \(Y\) and \(Z\)* is more reliable than extrapolation from \(Y\) alone
  - \(\text{sbj} \Rightarrow \text{salient (low confidence)} \Rightarrow \text{perspron}\)
  - \(\text{sbj and vf} \Rightarrow \text{salient (high confidence)} \Rightarrow \text{perspron}\)

\[ P(\text{perspron} \mid \text{sbj}, \text{vf}) \geq P(\text{perspron} \mid \text{sbj}) \]

* Given that \(Y_{\text{sal}}\) and \(Z_{\text{sal}}\) point to the same degree of salience
One or two dimensions?

- Multidimensional prediction

  - Prediction 1 may hold \( P(X_{\text{sal}} | Y_{\text{sal}}) > P(X_{\text{sal}}) \)
    - But only if \( X_{\text{sal}} \) and \( Y_{\text{sal}} \) are affected by the same dimension of salience

  - Prediction 2 does not hold \( P(X_{\text{sal}} | Y_{\text{sal}}, Z_{\text{sal}}) \geq P(X_{\text{sal}} | Z_{\text{sal}}) \)
    - If \( X_{\text{sal}} \) is determined by one dimension of salience and \( Y_{\text{sal}} \) by another dimension of salience
One or two dimensions?

Prediction 1

Probability increase confirmed
- if there are multiple dimensions of salience, they are interrelated

| realization $X_{sal\uparrow}$ | condition $Y_{sal\uparrow}$ | (conditioned) probability $P(X_{sal\uparrow}|Y_{sal\uparrow})$ | probability increase (vs. unconditioned) |
|-------------------------------|-----------------------------|-------------------------------------------------|--------------------------------------|
| perspron                      | (none)                      | 10.80% (8,557/79,222)                           | +0.63%                               |
|                               | vf                          | 11.43%                                          |                                      |
|                               | sbj                         | 20.06%                                          |                                      |
|                               | (none)                      | 42.50% (33,667/79,222)                          | +9.26%                               |
| sbj                           | perspron                    | 78.94%                                          | +36.44%                              |
|                               | vf                          | 63.91%                                          |                                      |
|                               | (none)                      | 33.16% (16,789/79,222)                          | +21.41%                              |
|                               | perspron                    | 35.08%                                          | +1.92%                               |
|                               | sbj                         | 49.87%                                          | +16.71%                              |

Significant positive correlation between perspron, sbj, vf

<table>
<thead>
<tr>
<th>realization</th>
<th>$\chi^2$</th>
<th>$\phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>±perspron</td>
<td>±vf</td>
<td>p &lt; .0001</td>
</tr>
<tr>
<td>±perspron</td>
<td>±sbj</td>
<td>p &lt; .0001</td>
</tr>
<tr>
<td>±sbj</td>
<td>±vf</td>
<td>p &lt; .0001</td>
</tr>
</tbody>
</table>
One or two dimensions?

Prediction 2

\[ P(X_{\text{sal}} | Y_{\text{sal}}, Z_{\text{sal}}) \geq P(X_{\text{sal}} | Z_{\text{sal}}) \]

- \( P(\text{perspron} | \text{vf, sbj}) < P(\text{perspron} | \text{sbj}) \)
- \( P(\text{vf} | \text{perspron, sbj}) < P(\text{vf} | \text{sbj}) \)
- **Direct counterevidence** for monodimensional models of salience
  - perspron is primarily determined by one dimension of salience
  - vf is primarily determined by another dimension of salience
  - sbj is sensitive to both dimensions
Forward-looking/backward-looking?

• Background
  ⇒ Salience influences information packaging
    pronominalization, subject role, sentence-initial position

• Corpus study 1
  ⇒ (at least) two dimensions of salience

• Corpus study 2
  Forward-looking vs. Backward-looking salience?

• Discussion
Forward-looking/backward-looking?

Multidimensional models of salience

„anaphoric“ (backward-looking)
„givenness“
„anaphora“ (attention guidance)

„cataphoric“ (forward-looking)
„emphasis“
„anadeixis“

(Givón 1983, 2001)
(Clamons et al. 1993, Mulkern 2007)
(Ehlich 1982, Cornish 2007)
Forward-looking/backward-looking?

Multidimensional models of salience

„anaphoric“ (backward-looking)
„anaphora“
„givenness“
„anaphora“

„cataphoric“ (forward-looking)
„emphasis“
„foregrounding“
„anadeixis“ (attention guidance)

Defined with respect to the preceding discourse / common ground
Attention-shifting operations / preparation for subsequent discourse

(Givón 1983, 2001)
(Clamons et al. 1993, Mulkern 2007)
(Ehlich 1982, Cornish 2007)
Forward-looking/backward-looking?

„backward-looking“

Covers most salience factors that are accessible to the hearer

Salience ~ attention:
Approximates attentional states of the hearer

Realization and distribution of the referent in previous discourse

Defined with respect to the preceding discourse / common ground

Attention-shifting operations / preparation for subsequent discourse

Generic labels

General characterization

Heuristic measurements

Functions
Forward-looking/backward-looking?

- "backward-looking"
  - Covers most salience factors that are accessible to the hearer
  - Salience ~ attention: Approximates attentional states of the hearer

- "forward-looking"
  - Includes sources of information that are available to the speaker only
  - For example, his/her intentions for the development of subsequent discourse

Can be partially reconstructed from:

- Realization and distribution of the referent in previous discourse
- Realization and distribution of the referent in subsequent discourse
- Defined with respect to the preceding discourse / common ground
- Attention-shifting operations / preparation for subsequent discourse

Generic labels
- General characterization
  - Heuristic measurements

Functions
Forward-looking/backward-looking?

- "backward-looking"
  - ±given
    - previous mention
  - Realization and distribution of the referent in previous discourse
  - Defined with respect to the preceding discourse / common ground
- "forward-looking"
  - ±important
    - subsequent mention
  - Realization and distribution of the referent in subsequent discourse
  - Attention-shifting operations / preparation for subsequent discourse

Robust, coarse-grained heuristic measurements

Abstract from theory-specific details

Different measurements with a variety of factors have been proposed
(cf. Chiarcos 2010 for an overview)
Forward-looking/backward-looking?

"backward-looking"  "forward-looking"

±given  previous mention  ±important  subsequent mention

Extrapolated from coreference annotation in TüBa-D/Z

Robust, coarse-grained heuristic measurements

<table>
<thead>
<tr>
<th>realization</th>
<th>±given</th>
<th>±important</th>
</tr>
</thead>
<tbody>
<tr>
<td>±perspron</td>
<td>$\chi^2$</td>
<td>.342</td>
</tr>
<tr>
<td>±sblj</td>
<td>$p &lt; .0001$</td>
<td>.288</td>
</tr>
<tr>
<td>±vf</td>
<td>$p &lt; .0001$</td>
<td>.065</td>
</tr>
</tbody>
</table>

Significant and positive correlation between heuristic measurements and packaging phenomena

But how do ±given and ±important interact?
Forward-looking/backward-looking?

• How do ±given and ±important interact?
• Experiment with C4.5 decision trees to predict packaging preferences from only ±given and ±important

(a) referring expressions  (b) grammatical roles  (c) word order

correctness: 34.6%  
(baseline: defNP, 33.6%)
correctness: 53.1%  
(baseline: sbj, 42.5%)
correctness: 38.7%  
(baseline: mf_initial, 33.6%)

+given  
|  +important: perspron
|  -important: defNP
-given: defNP

+given: sbj  
-given
|  +important: sbj
|  -important: other

+given: mf_initial  
-given
|  +important: vf
|  -important: mf_noninitial

• Important here is not the quality of the classification, but the predicted effects of ±given and ±important on information packaging
Packaging predictions

<table>
<thead>
<tr>
<th>+given</th>
<th>+important</th>
<th>-important</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal pronoun Subject</td>
<td>Mittelfeld initial</td>
<td>Mittelfeld initial</td>
</tr>
<tr>
<td>-given</td>
<td>Definite NP Subject</td>
<td>Mittelfeld non-initial</td>
</tr>
<tr>
<td>Definite NP Subject Vorfeld</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Definite NP Oblique Mittelfeld non-initial</td>
<td></td>
</tr>
</tbody>
</table>

This distribution explains the observations of first corpus study

- correlation between pronominalization and subject (+important, +given)
- correlation between vorfeld and subject (+important, -given)
- dispreference for subject pronouns (+given) in vorfeld (-given)

±given and ±important account for the observed distribution of grammatical devices
Discussion

- **Background**
  - salience influences information packaging

- **Corpus study 1**
  - (at least) two dimensions of salience

- **Corpus study 2**
  - these dimensions may be forward-looking and backward-looking salience
    - ±given and ±important account for the observed distribution
Results

- If a salience-based approach on information packaging is adopted to account for
  - the choice of referring expressions,
  - the assignment of grammatical roles, and
  - word order preferences in German,

- it is
  - necessary to distinguish (at least) two dimensions of salience in discourse, and
  - plausible to model these dimensions as backward-looking/hearer-oriented salience and forward-looking/speaker-oriented salience
Thank you
• Related research
  • Antecedent selection preferences of Finnish pronouns
    • Experimental support for a two- (or higher-) dimensional model of salience
  • German vorfeld
    • Empirical evidence and theoretical claims that the positioning in the vorfeld cannot explained solely on the basis of backward-looking salience/givenness
Related research: Pronouns

  - antecedent selection preferences for personal pronouns and demonstrative pronouns in Finnish
    - Personal pronoun more sensitive to grammatical role
    - Demonstrative pronoun more sensitive to word order
  \[ \Rightarrow \text{A unified notion of salience cannot be the sole determinant of the choice of referring expressions} \]

- But
  - constraints on the surface realization of antecedent-anaphor pairs are insufficient to disprove the existence of a unified \textit{cognitive} dimension of salience
    - see next slides for an alternative explanation
Related research: Pronouns

- An alternative explanation
  - one cognitive dimension of salience
  - salience-based grammaticalization
    - conventional associations between the linguistic realization of the antecedent and the referring expression of the anaphor
  - Pronominal anaphors with subject antecedent may evolve into syntactically bound pronouns
    - Cf. German (bound) relative pronoun *das* `that` from original (free) demonstrative pronoun
An alternative explanation

- One cognitive dimension of salience
- Salience-based grammaticalization
- Form-sensitive antecedent selection preferences for different types of pronouns may reflect different degrees of grammaticalization
  - Conventional associations may apply independently from the actual degree of salience a referent has
Related research: Pronouns

- An alternative explanation
  - one cognitive dimension of salience
  - salience-based **grammaticalization**
  - form-sensitive antecedent selection preferences for different types of pronouns may reflect different degrees of grammaticalization

⇒ Dimensionality of salience needs to be confirmed independently from the surface realization of the antecedent
  motivation for this study
Related research: *Vorfeld*

- **Word order in German**
  - „standard view“
    - *Vorfeld* marks topical (given) referents
    - Indefinite object tend to precede definite subjects in German OVS sentences
  - Speyer (2007)
    - 51% of *Vorfeld* constituents could neither semantically nor anaphorically linked to the preceding discourse
  - Dipper & Zinsmeister (2009)
    - 55% of *Vorfeld* constituents stand in no obvious relationship to the preceding discourse
Related research: Vorfeld

- **Word order in German**
    - Canonical topic position in German is the Wackernagel position (*Mittelfeld* initial)
    - Pragmatically-driven *Vorfeld* positioning (A’ movement) requires an additional pragmatic motivation
      - *kontrast* (Vallduví & Vilkuna 1998)
      - If the *Vorfeld* is not occupied by A’ movement, the highest-ranking *Mittelfeld* constituent is moved in the *Vorfeld* (formal movement)
        - this may be the topic
  
  \[\Rightarrow\text{Association between (givenness-)topic and Vorfeld is secondary}\]

  The primary function of the *vorfeld* is not to mark *givenness*
Related research: Vorfeld

- Alternative determinants of Vorfeld positioning in German
  - „discourse aboutness“ (Filippova & Strube 2007)
    Vorfeld constituents refer to the global discourse topic
    (= headline of a biographical article)
  - contrast & frame-setting topics (Speyer 2007)
    primary determinants of Vorfeld positioning
    backward-looking salience (Grosz et al. 1995) is secondary
Related research: Vorfeld

- Aboutness, contrast and frame-setting are speaker-oriented salience factors
  - speaker-private information (prior to utterance)
  - may belong to the same group of factors as ±important

⇒ Replace backward-looking /forward-looking dichotomy by hearer-oriented vs. speaker-oriented

forward-looking factors do, however, represent only a fraction of possible speaker-oriented salience factors

(Chiarcos 2010)
References


References


