Charting the Benefits of a New Perspective on Over-specification

Guanyi Chen[1], Kees van Deemter[12] and Chenghua Lin[2]

[1]Information and Computing Sciences, Utrecht University
[2]Computing Science, University of Aberdeen

CoMPPrag 2018
September 27th, 2018

Overview

(Re-)Defining Over-specification

MTuna Corpora and Findings

Future Work

Overview

(Re-)Defining Over-specification
- Gricean Maxim of Quantity
- TYPE in Referring Expressions

MTuna Corpora and Findings

Future Work

Overview

Gricean Maxim of Quantity [Grice, 1975]

The speaker should:

1. Include Enough information to allow an addressee to identify an intended referent;
2. Not be more informative than necessary.

The first rule defines the concept of an *distinguishing description*: a description $\mathcal{D}$ should be able to single out the referent r from distractors, i.e., $\bigcap_{P_i \in \mathcal{D}} [\![P_i]\!] = \{r\}$.

Gricean Maxim of Quantity [Grice, 1975]

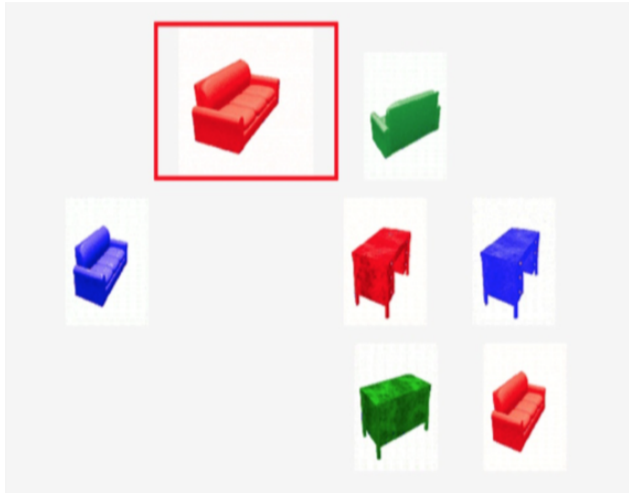
The speaker should:

1. Include Enough information to allow an addressee to identify an intended referent;
2. Not be more informative than necessary.

The first rule defines the concept of an *distinguishing description*: a description $\mathcal{D}$ should be able to single out the referent r from distractors, i.e., $\bigcap_{P_i \in \mathcal{D}} [\![P_i]\!] = \{r\}$.

Tuna Corpora [Deemter et al., 2012]

- Focusing on an assessment of the *humanlikeness* of the logical forms (do not rely on linguistic form) generated by a given REG algorithm;
- Evaluated by DICE [Dice, 1945]:



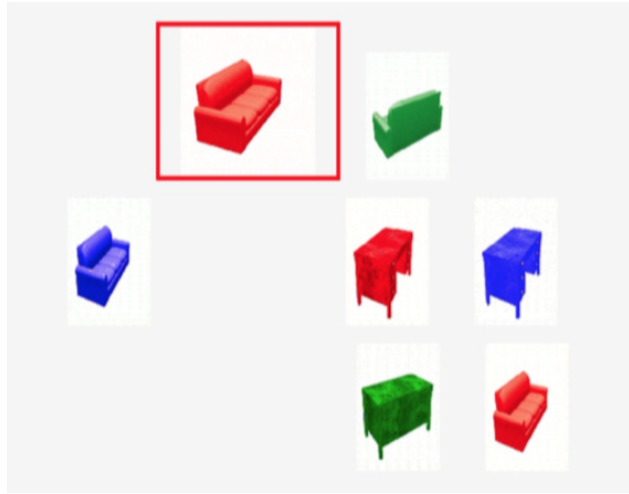
$$DICE(\mathcal{D}_H, \mathcal{D}_A) = \frac{2 \times |\mathcal{D}_H \cap \mathcal{D}_A|}{|\mathcal{D}_H| + |\mathcal{D}_A|}$$

where $\mathcal{D}_* = \{P_1, ..., P_n\}$ ($\{\cdot\}$ is a bag).

- Furniture corpus (simple) vs. People corpus (hard).
- Referring to a single object vs. Referring a set of two objects.
- A Mandarin version: MTuna
- High DICE score $\not\Rightarrow$ a distinguishing description.

Tuna Corpora [Deemter et al., 2012]

- Focusing on an assessment of the *humanlikeness* of the logical forms (do not rely on linguistic form) generated by a given REG algorithm;
- Evaluated by DICE [Dice, 1945]:



$$DICE(\mathcal{D}_H, \mathcal{D}_A) = \frac{2 \times |\mathcal{D}_H \cap \mathcal{D}_A|}{|\mathcal{D}_H| + |\mathcal{D}_A|}$$

where $\mathcal{D}_* = \{P_1, ..., P_n\}$ ($\{\cdot\}$ is a bag).

- Furniture corpus (simple) vs. People corpus (hard).
- Referring to a single object vs. Referring a set of two objects.
- A Mandarin version: MTuna
- High DICE score $\not\Rightarrow$ a distinguishing description.

A Definition of Over-specification

[Engelhardt et al., 2006, Koolen et al., 2011, Engelhardt et al., 2011] called a RE over-specified if it breaks the second rule of Gricean Maxim of Quantity.

- BUT, how to define the situation of a RE with only necessary information (*Minimal Description*)?
- One definition that is often used: None of the properties in D can be removed, i.e., $\nexists P(P \in \mathcal{D} \wedge \bigcap_{P_i \in \mathcal{D}-\{P\}} [\![P_i]\!] = \{r\})$

A Definition of Over-specification

[Engelhardt et al., 2006, Koolen et al., 2011, Engelhardt et al., 2011] called a RE over-specified if it breaks the second rule of Gricean Maxim of Quantity.

- BUT, how to define the situation of a RE with only necessary information (*Minimal Description*)?
- One definition that is often used: None of the properties in D can be removed, i.e., $\nexists P(P \in \mathcal{D} \wedge \bigcap_{P_i \in \mathcal{D}-\{P\}} [\![P_i]\!] = \{r\})$

A Definition of Over-specification

[Engelhardt et al., 2006, Koolen et al., 2011, Engelhardt et al., 2011] called a RE over-specified if it breaks the second rule of Gricean Maxim of Quantity.

- BUT, how to define the situation of a RE with only necessary information (*Minimal Description*)?
- One definition that is often used: None of the properties in D can be removed, i.e., $\nexists P(P \in \mathcal{D} \wedge \bigcap_{P_i \in \mathcal{D}-\{P\}} [\![P_i]\!] = \{r\})$

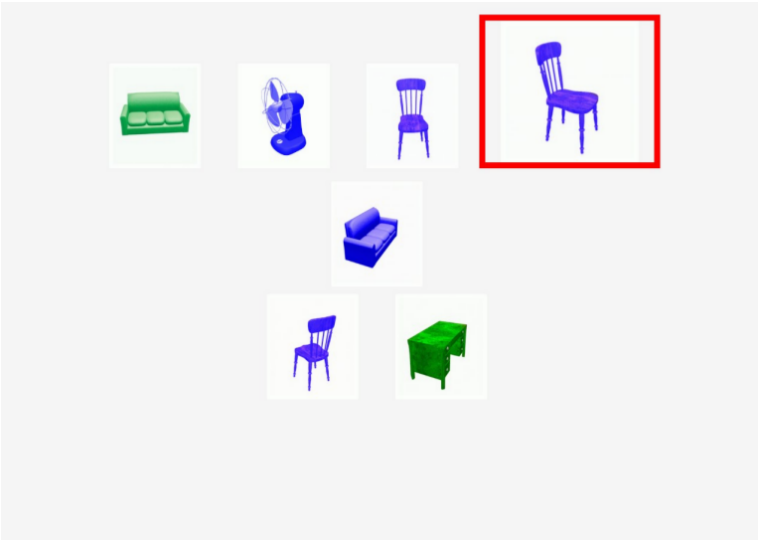
A Definition of Over-specification

[Engelhardt et al., 2006, Koolen et al., 2011, Engelhardt et al., 2011] called a RE over-specified if it breaks the second rule of Gricean Maxim of Quantity.

- BUT, how to define the situation of a RE with only necessary information (*Minimal Description*)?
- One definition that is often used: None of the properties in D can be removed, i.e., $\nexists P(P \in \mathcal{D} \wedge \bigcap_{P_i \in \mathcal{D}-\{P\}} [\![P_i]\!] = \{r\})$

$\mathcal{D} = \{\texttt{COLOUR} = \textit{blue}, \texttt{SIZE} = \textit{small}\}$

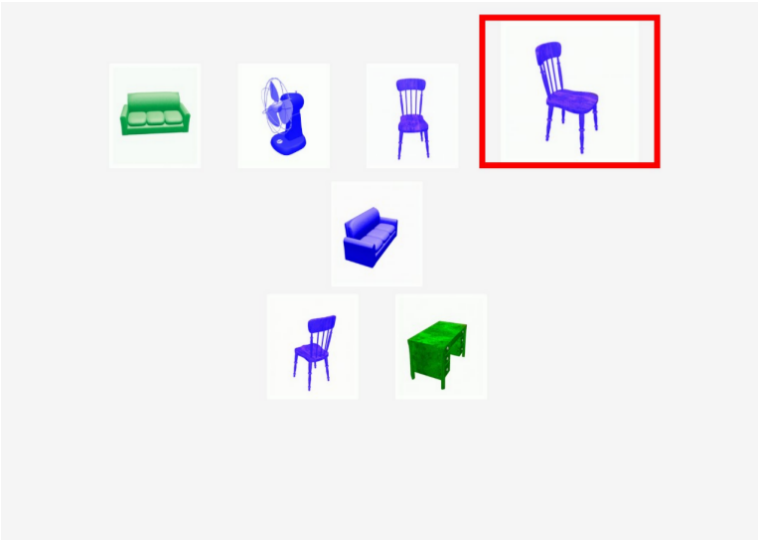
An Example



- $\mathcal{D}_1 = \{\texttt{SIZE} = \textit{large}\}$
- $\mathcal{D}_2 = \{\texttt{ORIENTATION} = \textit{right}, \texttt{TYPE} = \textit{chair}\}$
- $|\mathcal{D}_2| > |\mathcal{D}_1|$
- $\nexists P(P \in \mathcal{D}_2 \wedge \bigcap_{P_i \in \mathcal{D}_2 - \{P\}} [\![P_i]\!] = \{r\})$

- A broader definition of *Minimal Description* [Dale and Reiter, 1995]: a RE $\mathcal{D} = \{P_1, ..., P_n\}$, where there is no distinguishing description $\mathcal{D}' = \{P_1, ..., P_m\}$ such that $m < n$ (that is, $|\mathcal{D}'| < |\mathcal{D}|$);
- $\mathcal{D}_2 :=$ *Numerical Over-specification*.
- Can numerical over-specification help listeners to identify targets? Or NOT? Or the OPPOSITE?

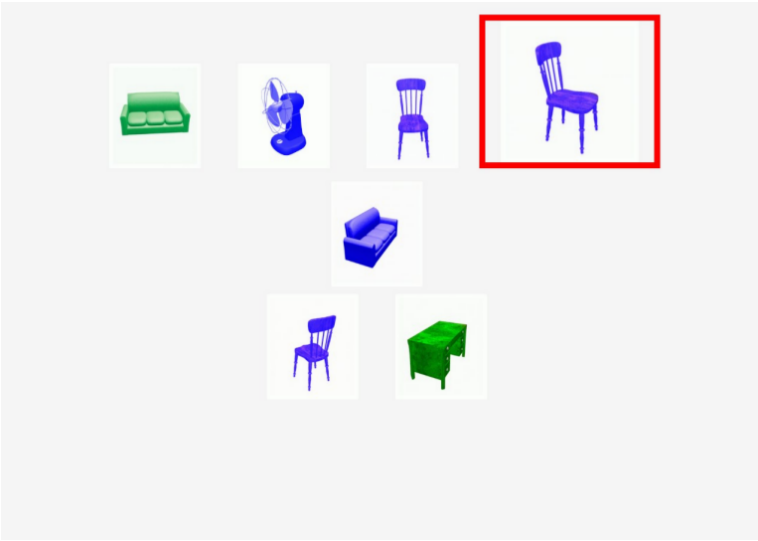
An Example



- $\mathcal{D}_1 = \{\texttt{SIZE} = \textit{large}\}$
- $\mathcal{D}_2 = \{\texttt{ORIENTATION} = \textit{right}, \texttt{TYPE} = \textit{chair}\}$
- $|\mathcal{D}_2| > |\mathcal{D}_1|$
- $\nexists P(P \in \mathcal{D}_2 \wedge \bigcap_{P_i \in \mathcal{D}_2 - \{P\}} [\![P_i]\!] = \{r\})$

- A broader definition of *Minimal Description* [Dale and Reiter, 1995]: a RE $\mathcal{D} = \{P_1, ..., P_n\}$, where there is no distinguishing description $\mathcal{D}' = \{P_1, ..., P_m\}$ such that $m < n$ (that is, $|\mathcal{D}'| < |\mathcal{D}|$);
- $\mathcal{D}_2 :=$ *Numerical Over-specification*.
- Can numerical over-specification help listeners to identify targets? Or NOT? Or the OPPOSITE?

An Example



- $\mathcal{D}_1 = \{\texttt{SIZE} = \textit{large}\}$
- $\mathcal{D}_2 = \{\texttt{ORIENTATION} = \textit{right}, \texttt{TYPE} = \textit{chair}\}$
- $|\mathcal{D}_2| > |\mathcal{D}_1|$
- $\nexists P(P \in \mathcal{D}_2 \wedge \bigcap_{P_i \in \mathcal{D}_2 - \{P\}} [\![P_i]\!] = \{r\})$

- A broader definition of *Minimal Description* [Dale and Reiter, 1995]: a RE $\mathcal{D} = \{P_1, ..., P_n\}$, where there is no distinguishing description $\mathcal{D}' = \{P_1, ..., P_m\}$ such that $m < n$ (that is, $|\mathcal{D}'| < |\mathcal{D}|$);
- $\mathcal{D}_2 :=$ *Numerical Over-specification*.
- Can numerical over-specification help listeners to identify targets? Or NOT? Or the OPPOSITE?

Hypotheses (for both MTuna and ETuna)

- More over-specifications in People (harder) corpus;
- More under-specifications in People (harder) corpus;
- Numerical Over-specifications occurs in the corpus.

Overview

(Re-)Defining Over-specification

- Gricean Maxim of Quantity
- TYPE in Referring Expressions

MTuna Corpora and Findings

Future Work

Results in MTuna (single referent)

	#RE	Over-specifications
Furniture	224	166
People	256	200

Results in MTuna (single referent)

	#RE	Over-specifications	
Furniture	224	166	
People	256	200	($p > .1$)

Results in MTuna (single referent)

	#RE	Over-specifications	
Furniture	224	166	
People	256	200	($p > .1$)

(1) a. (MD) 红色 的 / the red object
b. 红色 的 桌子 / the red table

Results in MTuna (single referent)

	#RE	Over-specifications	
Furniture	224	166	
People	256	200	($p > .1$)

	#RE	**Real** Over-specifications	Nominal Over-specification
Furniture	224	83	83
People	256	146	54

Formal Definitions

1. *Real Over-specification* arises when a description has superfluous non-TYPE attributes, i.e., a description $\mathcal{D} = \{P_1, ..., P_n\}$ where at least one of the $P \epsilon \mathcal{D}$ is such that $P \neq \texttt{TYPE}$ and $\bigcap_{P_j \in \mathcal{D} - \{P\}} [\![P_j]\!] = \{r\}$.
2. *Nominal Over-specification* is a description $\mathcal{D}$ in which any $P \in \mathcal{D}$ that causes $\bigcap_{P_j \in \mathcal{D} - \{P\}} [\![P_j]\!] = \{r\}$ is TYPE; in other words, only TYPE attributes are superfluous, no other attributes is superfluous.

Other Issues Related to TYPE

- [Dale and Reiter, 1995] (IA) added a provision to ensure that each logical form generated contains a TYPE (to ensure REs have head nouns);
- BUT, it not always true for some languages or domains.
- There are 97% and 85% superfluous TYPE attributes in English and Chinese, respectively;
- In MTuna, there are much more superfluous TYPE in furniture corpus (94%) than people corpus (74%).
  1. People corpus has only one type of TYPE: person;
  2. Furniture corpus has four types of TYPE: chair, fan, sofa and table.

Other Issues Related to TYPE

- [Dale and Reiter, 1995] (IA) added a provision to ensure that each logical form generated contains a TYPE (to ensure REs have head nouns);
- BUT, it not always true for some languages or domains.
- There are 97% and 85% superfluous TYPE attributes in English and Chinese, respectively;
- In MTuna, there are much more superfluous TYPE in furniture corpus (94%) than people corpus (74%).
  1. People corpus has only one type of TYPE: person;
  2. Furniture corpus has four types of TYPE: chair, fan, sofa and table.

Other Issues Related to TYPE

- [Dale and Reiter, 1995] (IA) added a provision to ensure that each logical form generated contains a TYPE (to ensure REs have head nouns);
- BUT, it not always true for some languages or domains.
- There are 97% and 85% superfluous TYPE attributes in English and Chinese, respectively;
- In MTuna, there are much more superfluous TYPE in furniture corpus (94%) than people corpus (74%).
  1. People corpus has only one type of TYPE: person;
  2. Furniture corpus has four types of TYPE: chair, fan, sofa and table.

Overview

(Re-)Defining Over-specification

MTuna Corpora and Findings

Future Work

MTuna corpora

- Mandarin Chinese version of the Tuna corpora [van Deemter et al., 2017b];
- Most trials are inherited from Tuna experiment, but it also has extras;
- Two settings: REs in subject positions or in object positions [van Deemter et al., 2017a]:
  1. ____在红色方块中/____zai hongse fangkuai zhong
  2. 红色方块中的是____/hongse fangkuai zhong de shi ____
- In some trials in MTuna, `TYPE` is used for distinguishing objects.

Some Results in MTuna[1]

	total	minimal	real	nom	num	under
Furniture	399	46	135	128	24	66
People	400	17	246	68	14	54

1. No significant difference between the proportion of under-specifications in Furnture and People corpus ($p > .1$);
2. More real over-specifications and fewer minimal descriptions in the people corpus ($p < .01$);
3. 5% of REs were numerical over-specifications;
4. There are more over-specifications and fewer under-specifications in subject position ($p < .01$).

[1] Exclude REs that use location and REs that refer to the wrong object.

Some Results in MTuna[1]

	total	minimal	real	nom	num	under
Furniture	399	46	135	128	24	66
People	400	17	246	68	14	54

1. No significant difference between the proportion of under-specifications in Furnture and People corpus ($p > .1$);
2. More real over-specifications and fewer minimal descriptions in the people corpus ($p < .01$);
3. 5% of REs were numerical over-specifications;
4. There are more over-specifications and fewer under-specifications in subject position ($p < .01$).

[1] Exclude REs that use location and REs that refer to the wrong object.

Some Results in MTuna[1]

	total	minimal	real	nom	num	under
Furniture	399	46	135	128	24	66
People	400	17	246	68	14	54

1. No significant difference between the proportion of under-specifications in Furnture and People corpus ($p > .1$);
2. More real over-specifications and fewer minimal descriptions in the people corpus ($p < .01$);
3. 5% of REs were numerical over-specifications;
4. There are more over-specifications and fewer under-specifications in subject position ($p < .01$).

[1] Exclude REs that use location and REs that refer to the wrong object.

Some Results in MTuna[1]

	total	minimal	real	nom	num	under
Furniture	399	46	135	128	24	66
People	400	17	246	68	14	54

1. No significant difference between the proportion of under-specifications in Furnture and People corpus ($p > .1$);
2. More real over-specifications and fewer minimal descriptions in the people corpus ($p < .01$);
3. 5% of REs were numerical over-specifications;
4. There are more over-specifications and fewer under-specifications in subject position ($p < .01$).

[1] Exclude REs that use location and REs that refer to the wrong object.

Some Results in MTuna[1]

	total	minimal	real	nom	num	under
Furniture	399	46	135	128	24	66
People	400	17	246	68	14	54

1. No significant difference between the proportion of under-specifications in Furnture and People corpus ($p > .1$);
2. More real over-specifications and fewer minimal descriptions in the people corpus ($p < .01$);
3. 5% of REs were numerical over-specifications;
4. There are more over-specifications and fewer under-specifications in subject position ($p < .01$).

[1] Exclude REs that use location and REs that refer to the wrong object.

Summary so far

- Over- and under-specification: the standard view;
- A new perspective on specification
- Using this perspective to understand REs in a corpus

Overview

(Re-)Defining Over-specification

MTuna Corpora and Findings

Future Work

Referring to Plural Referents (a set)

(2) a. the red table and the red chair
b. the red table and chair
c. the red furniture

- From 2a to 2b is syntactic aggregation;
- From 2b to 2c is semantic aggregation.

Comparing between Languages (MTuna vs. (E)Tuna)

	minimal	real	nom	num	under
Furniture (Mandarin)	3.6	37.1	37.1	0	18.6
People (Mandarin)	6.3	57	21.1	0.4	14.5
Furniture (English)	0.7	41.9	37.5	0	19.6
People (English)	2.6	77.9	7.1	0	12.3

Table: Results in overlapped singular portion (normalised).

1. No significant difference between under-spec in two Languages ($p > .1$), much more than expected;
2. More real over-specifications and less minimal descriptions in ETuna ($p < .01$) (*man* vs. 男人);
3. Numerical over-specification never appears in ETuna;
4. ETuna has more superfluous `TYPE` attributes ($> 97\%$) than that in MTuna ($p < .01$);
5. More analysis to come...

Comparing between Languages (MTuna vs. (E)Tuna)

	minimal	real	nom	num	under
Furniture (Mandarin)	3.6	37.1	37.1	0	18.6
People (Mandarin)	6.3	57	21.1	0.4	14.5
Furniture (English)	0.7	41.9	37.5	0	19.6
People (English)	2.6	77.9	7.1	0	12.3

Table: Results in overlapped singular portion (normalised).

1. No significant difference between under-spec in two Languages ($p > .1$), much more than expected;
2. More real over-specifications and less minimal descriptions in ETuna ($p < .01$) (*man* vs. 男人);
3. Numerical over-specification never appears in ETuna;
4. ETuna has more superfluous `TYPE` attributes ($> 97\%$) than that in MTuna ($p < .01$);
5. More analysis to come...

Comparing between Languages (MTuna vs. (E)Tuna)

	minimal	real	nom	num	under
Furniture (Mandarin)	3.6	37.1	37.1	0	18.6
People (Mandarin)	6.3	57	21.1	0.4	14.5
Furniture (English)	0.7	41.9	37.5	0	19.6
People (English)	2.6	77.9	7.1	0	12.3

Table: Results in overlapped singular portion (normalised).

1. No significant difference between under-spec in two Languages ($p > .1$), much more than expected;
2. More real over-specifications and less minimal descriptions in ETuna ($p < .01$) (*man* vs. 男人);
3. Numerical over-specification never appears in ETuna;
4. ETuna has more superfluous TYPE attributes ($> 97\%$) than that in MTuna ($p < .01$);
5. More analysis to come...

Comparing between Languages (MTuna vs. (E)Tuna)

	minimal	real	nom	num	under
Furniture (Mandarin)	3.6	37.1	37.1	0	18.6
People (Mandarin)	6.3	57	21.1	0.4	14.5
Furniture (English)	0.7	41.9	37.5	0	19.6
People (English)	2.6	77.9	7.1	0	12.3

Table: Results in overlapped singular portion (normalised).

1. No significant difference between under-spec in two Languages ($p > .1$), much more than expected;
2. More real over-specifications and less minimal descriptions in ETuna ($p < .01$) (*man* vs. 男人);
3. Numerical over-specification never appears in ETuna;
4. ETuna has more superfluous `TYPE` attributes ($> 97\%$) than that in MTuna ($p < .01$);
5. More analysis to come...

Conclusions

1. More real over-specifications in people (harder) domain;
2. Use of `TYPE` depends on languages and domains;
3. More under-specifications than expected (5%),
   Should REG algorithms sometimes underspecify as well?
4. Over-specification does not always involve a superfluous property;
   studies of over-specification should include numerical
   over-specification.

Conclusions

1. More real over-specifications in people (harder) domain;
2. Use of `TYPE` depends on languages and domains;
3. More under-specifications than expected (5%),
   Should REG algorithms sometimes underspecify as well?
4. Over-specification does not always involve a superfluous property;
   studies of over-specification should include numerical
   over-specification.

Many Thanks!

References I


Dale, R. and Reiter, E. (1995).
Computational interpretations of the gricean maxims in the generation of referring expressions.
Cognitive science, 19(2):233–263.


Deemter, K. v., Gatt, A., Sluis, I. v. d., and Power, R. (2012).
Generation of referring expressions: Assessing the incremental algorithm.
Cognitive science, 36(5):799–836.


Dice, L. R. (1945).
Measures of the amount of ecologic association between species.
Ecology, 26(3):297–302.

References II


Engelhardt, P. E., Bailey, K. G., and Ferreira, F. (2006).
Do speakers and listeners observe the gricean maxim of quantity?
Journal of Memory and Language, 54(4):554–573.


Engelhardt, P. E., Demiral, Ş. B., and Ferreira, F. (2011).
Over-specified referring expressions impair comprehension: An erp study.
Brain and cognition, 77(2):304–314.


Grice, H. P. (1975).
Logic and conversation.
1975, pages 41–58.


Koolen, R., Gatt, A., Goudbeek, M., and Krahmer, E. (2011).
Factors causing overspecification in definite descriptions.
Journal of Pragmatics, 43(13):3231–3250.

References III


van Deemter, K., Sun, L., and Sybesma, R. (2017a).
Definiteness: Towards a global perspective.


van Deemter, K., Sun, L., Sybesma, R., Li, X., Bo, C., and Yang, M. (2017b).
Investigating the content and form of referring expressions in mandarin: introducing the mtuna corpus.
In *Proceedings of the 10th International Conference on Natural Language Generation*, pages 213–217.