



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

Computerlinguistik

Lehreinheit 3-4: Multilinguale Mensch-Maschine
Kommunikation: Linguistische Aspekte und
Anwendungen-1

Dr. Christina Alexandris
Nationale Universität Athen
Deutsche Sprache und Literatur

Anwendungen der Semantik

Wie „versteht“ der Computer die Sprache?

Anwendungen der Semantik- Einführung

- In den meisten Systemen der Verarbeitung der natürlichen Sprache folgt auf die Analyse der syntaktischen Struktur der Eingaben die semantische Analyse.
- Das bedeutet, dass der Satz als Eingabe natürlicher Sprache zuerst auf der **lexikalischen Ebene** mit Hilfe eines **Taggers**, und auf der **morphosyntaktischen Ebene** mit Hilfe eines **Parsers** analysiert wird.
- Anschließend folgt die **semantische** Analyse (Schneider, 2002). Dabei wird die semantische Struktur des Satzes analysiert .
- Allerdings kann in vielen Systemen die Analyse auf der morphosyntaktischen und auf der semantischen Ebene gleichzeitig stattfinden.



Im Gegensatz zur Syntax,

- die als die formale (nicht inhaltliche) Struktur der Ausdrücke einer Sprache definiert werden kann, bezieht sich die Semantik auf die inhaltliche (Bedeutungs-) Struktur der Ausdrücke einer Sprache (Schneider, 2002).
- Für die Analyse der semantischen Ebene von Eingaben natürlicher Sprache (z.B. Wörter, Phrasen und Sätze) werden **verschiedene theoretische Modelle und Verarbeitungsmethoden** verwendet.
- Diese Ansätze, die hier nicht beschrieben werden, reichen von relativ einfach bis sehr komplex.



Die Bedeutung natürlichsprachlicher Ausdrücke (Wörter, Sätze oder Texte)

- wird von der **Computersemantik** in entsprechende Algorithmen und Programme umgesetzt, damit sie von dem System "verstanden,, werden können (**maschinelle Bedeutungsbestimmung**) (Schielen und Klabunde, 2004).
- Ein bedeutender Teil der Computersemantik basiert auf den traditionellen, formalen Semantiktheorien und beschäftigt sich unter anderem auch mit der Analyse der sprachlichen **Ambiguitäten** die auf verschiedenen Ebenen von Sprache zu finden sind.



Ebenen sprachlicher Ambiguitäten

Ebenen sprachlicher **Ambiguitäten** (Schielen und Klabunde, 2004) sind:

- (a) lexikalische Ambiguitäten,
- (b) syntaktische Ambiguitäten und
- (c) semantische Ambiguitäten.



Beispiele

Beispiele (Schielen und Klabunde, 2004):

(a) Auf meinem Schreibtisch ist eine Maus!

(b) Die Touristen sahen die Störche auf ihrem
Flug nach Afrika.

(c) Ein Buch hat jeder gelesen.



Anwendungen und Probleme

Multilinguale Mensch-Maschine
Kommunikation: Linguistische Aspekte
und Anwendungen

Unifikationsgrammatiken

- Auf der syntaktischen Ebene wird versucht lexikalische, syntaktische und semantische Ambiguitäten mit Hilfe von Unifikationsgrammatiken zu reduzieren (Wasow et al, 2005).
- Unifikationsgrammatiken sind Grammatiktypen, in denen Merkmalstrukturen eine zentrale Rolle spielen, d.h. die syntaktischen Kategorien werden nicht als einzelne (atomare) Symbole konzipiert, sondern als **Merkmalstrukturen**.
- Diese Merkmalsstrukturen werden mit Hilfe der "**Unifikationsoperation**" miteinander verknüpft.



Unifikationsgrammatiken- Merkmalstrukturen

- Wenn z.B. zwei grammatische Kategorien den **gleichen Wert** für die **Merkmalstrukturen** Kasus, Numerus und Genus haben, dann sind sie **unifizierbar** und können kombiniert werden, um eine Konstituente wie, zum Beispiel, eine Phrase, zu bilden.
- Oft werden diese Merkmalstrukturen nicht in einer analytischen Form, sondern in einer generellen aber zugleich kompakten Form repräsentiert. Diese Repräsentationsform wird als "Unterspezifikation der **Merkmalstrukturen**" bezeichnet.



Unifikationsgrammatiken

Die wichtigsten Vertreter der Unifikationsgrammatiken sind:

(1) die **Generalized Phrase Structure Grammar (GPSG**, Gazdar et al., 1985),

(2) die **Lexical Functional Grammar (LFG**, Bresnan, 1982),

(3) **PATRII** (Shieber, 1986),

(4) die **Head-Driven Phrase Structure Grammar (HPSG**, Pollard und Sag, 1994).



Head-Driven Phrase Structure Grammar-HPSG

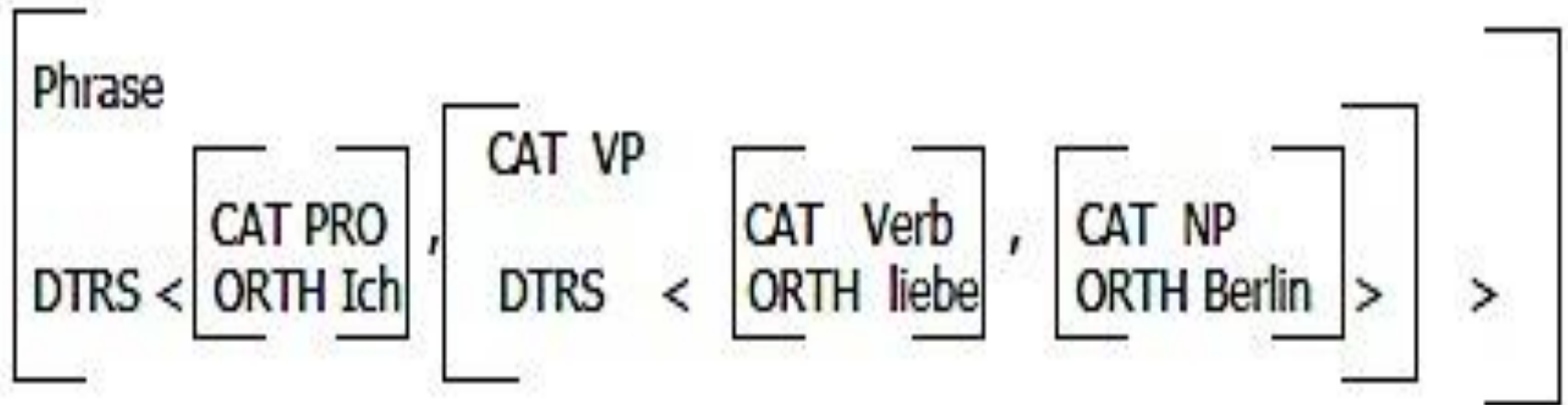
In der Head-Driven Phrase Structure Grammar (HPSG) werden sowohl:

- die syntaktischen als auch
- die semantischen Eigenschaften eines Wortes, bzw. lexikalischen Zeichens durch die Merkmalstruktur SYNSEM repräsentiert.
- Die transitiven Verben, die eine Subkategorisierung (wie Objekte, Subjekt) haben, enthalten im Gegensatz zu den intransitiven Verben die Merkmalstruktur SUBCAT.



Beispiel

(Struktur aus Jurafsky and Martin, 2008).



PRO = Pronomen

VP = Verbalphrase

NP = Nominalphrase



LFG Grammatiken

Im Gegensatz zu HPSG (und GPSG und PATR II), in der es nur eine Beschreibungsebene gibt, enthält die LFG Grammatik zwei Beschreibungsebenen:

(1) die C-Struktur -im Wesentlichen eine **Konstituentenstruktur**, die durch grammatische Regeln (eine kontextfreie Grammatik) beschrieben wird, (dabei werden auch Elemente des X' Modells verwendet) und

(2) die F-Struktur, die **syntaktische Funktionen**, wie Subjekt, Prädikat und Attribut enthält.



Einführung- Voraussetzungen

Maschinelle Übersetzung

Voraussetzungen für die maschinelle Übersetzung-1

Grundvoraussetzung für die maschinelle Übersetzung eines geschriebenen oder eines gesprochenen Textes ist, dass es in das System in **geschriebener** (elektronischer) Form eingegeben wird.

Das bedeutet, dass Texte, die in gedruckter Form erscheinen, wie ein Text einer Tageszeitung, zuerst in eine **geschriebene** (elektronische) Form umgewandelt werden sollen, zum Beispiel mit Hilfe eines Scanners.

Die Texte gesprochener Sprache werden von speziellen Systemen **Spracherkennung** (ASR, "Speech Recognition" systems) bearbeitet, durch die Signale der gesprochenen Sprache als Phoneme erkannt und mit Hilfe phonologischer Regeln in Buchstaben (Graphemen) und schließlich in Wörter eines geschriebenen Textes umgeformt werden.



Voraussetzungen für die maschinelle Übersetzung-2

Eine weitere Grundvoraussetzung ist, dass der Text, der maschinell übersetzt werden soll, möglichst "sauber" sein muss.

Das bedeutet, dass der Text keine Fehler bezüglich der Orthographie oder Interpunktion enthalten soll.

Der Text muss also für den Vorgang der maschinellen Übersetzung vorbereitet sein.

Die automatische Vorbereitung des Textes kann von Pre-Editing Systemen durchgeführt werden.



Weitere Voraussetzungen für die maschinelle Übersetzung-1

eines geschriebenen oder eines gesprochenen Textes hängen von der Art der maschinellen Übersetzung ab.

- Der ideale Text für eine gelungene maschinelle Übersetzung soll **einfache Formen** und **eindeutige Strukturen** bezüglich der morphosyntaktischen Ebene haben, er soll durch eine **deutliche pragmatische Struktur** charakterisiert sein und sich bezüglich der Semantik auf einen **bestimmten** Bereich (Domäne) bzw. "Welt" beschränken.
- Oft werden Texte (Korpora) aus einem bestimmten Bereich (Domäne) verwendet, in denen das System maschineller Übersetzung "arbeiten" soll.



Weitere Voraussetzungen für die maschinelle Übersetzung-2

- Diese idealen Bedingungen werden in den etwas strengeren Voraussetzungen für die maschinelle Übersetzung von v. Hahn (2001) ausgedrückt, nach denen der Text "schriftlich, domänenabhängig, einfach, stark strukturiert ist und wenn man dazu einen Korpus hat,, (v. Hahn, 2001).
- Viele Systeme maschineller Übersetzung können auch Texte behandeln, die nicht als "ideale Texte" bezeichnet werden und nicht allen dieser vorher beschriebenen Voraussetzungen entsprechen.



Voraussetzungen für die maschinelle Übersetzung

Maschinell kann man übersetzen, wenn der Text
schriftlich
domänenabhängig
einfach
stark strukturiert ist
und wenn man dazu einen Korpus hat

(v. Hahn, 2001)



Grundstruktur eines Systems maschineller Übersetzung-1

- In einem System maschineller Übersetzung wird der Ausgangssprachliche Text (Quellsprache) in kleineren Einheiten analysiert (**Analyse**), die mit äquivalenten Einheiten der Zielsprache (mapping) verknüpft werden (**Transfer**).
- Anschließend wird aus den Einheiten der Zielsprache ein Zielsprachlicher Text konstruiert und erzeugt (**Generierung**)
- Die Vorgänge der Analyse, des Transfers und der Generierung bilden die drei **Grundphasen** in jedem automatischen Übersetzungsprozess und sind in der Grundstruktur jedes Systems maschineller Übersetzung zu finden.
- Die Analyse des ausgangsprachlichen Texts in **kleineren Einheiten** betrifft vor allem die morphosyntaktische Ebene der Ausgangssprache in sehr vielen Systemen wird aber auch die semantische Ebene im Zusammenhang mit der morphosyntaktischen Ebene behandelt.



Grundstruktur eines Systems maschineller Übersetzung-2

- In manchen Systemen wird der Ausgangssprachliche Text auch in **pragmatischen Einheiten** analysiert.
- Der **Grad der Analyse** sowie die Art und Weise der Analyse hängt von der Struktur des Systems und von dem Ansatz (d.h. der Methode) ab, der für die maschinelle Übersetzung benutzt wird.
- In der Phase des **Transfers** werden die morphosyntaktischen und semantischen (eventuell auch die pragmatischen) Einheiten der Ausgangsprache mit den äquivalenten Einheiten der Zielsprache verknüpft (mapping).
- Aus den morphosyntaktischen und semantischen (oder auch pragmatischen) Einheiten der Zielsprache werden in der **Generierungsphase** Sätze oder auch Texte gebildet.
- Der von dem System generierte Text kann anschließend **vom Übersetzer verbessert** oder von einem **Post-Editing System** automatisch korrigiert werden.



Systeme maschineller Übersetzung- Aufteilungen-1

Systeme maschineller und computergestützter Übersetzung

- **MT = Full Machine Translation**

⇒ Der ganze Übersetzungsprozess wird völlig und nur von einem System maschineller Übersetzung durchgeführt

- **MAT = Machine Aided Translation**

⇒ Der Übersetzungsprozess wird teilweise von einem System maschineller Übersetzung durchgeführt

- HAMT = Human-Aided Machine Translation (überwiegend von System maschineller Übersetzung durchgeführt)
- MAMT = Machine-Aided Human Translation (überwiegend von Übersetzer durchgeführt, das System spielt eine helfende Rolle)



Systeme maschineller Übersetzung- Aufteilungen-2

Weitere Aufteilungen:

- (1) **Bilinguale** Systeme maschineller Übersetzung :
 - Unidirektionale Systeme maschineller Übersetzung
 - Bidirektionale Systeme maschineller Übersetzung
(Übersetzung ist für beide Richtungen vorgesehen)

Multilinguale Systeme maschineller Übersetzung

- (2) **Batch-Systeme** (**keine Interaktion** während der Übersetzung, Vor- und Nachbearbeitung vorgesehen)
Interaktive Systeme (erlauben die interaktive Auflösung der Ambiguitäten/ eine Auswahl von Übersetzungsalternativen)

Ein interaktives System kann Vorschläge machen/stellt Fragen an den Benutzer, um so das dem System fehlende Weltwissen zu kompensieren.



Divergenzen Lexikalische Lücken

Übersetzungsschwierigkeiten

Übersetzungsschwierigkeiten

Zu den wichtigsten Kategorien der Übersetzungsschwierigkeiten gehören:

- die **Divergenzen**, die als Unterschiede zwischen zwei Sprachen in den syntaktischen Strukturen definiert werden können (Dorna und Jekat, 2004) und
- die **Lexikalischen Lücken** (gaps) oder **Nichtentsprechungen** (mismatches), die als versprachlichte Konzepte definiert werden können, die nicht in der gleichen Form in zwei Sprachen existieren (Dorna und Jekat, 2004).



Divergenzen

- Das Partizip "auszutauschende" in der deutschen Phrase "Der auszutauschende Filter muss gesondert entsorgt werden" kann im Griechischen in verschiedenen Weisen anerkannt und verarbeitet werden.
- Somit besteht die Möglichkeit, dass zwei oder mehrere äquivalente Sätze erzeugt werden können, wie zum Beispiel der Satz: "Το προς αντικατάσταση φίλτρο πρέπει να πεταχτεί ξεχωριστά" und der Satz: "Το φίλτρο που πρέπει να αντικατασταθεί πρέπει να πεταχτεί ξεχωριστά".
- Ein **weiteres Beispiel** ist der Satz: "Ein unter Druck geratenes Ventil nicht öffnen", der im Griechischen sowohl mit dem Satz "Μην ανοίγετε μια υπό πίεση ευρισκομένη βαλβίδα" als auch mit dem Satz "Μην ανοίγετε μια βαλβίδα που βρίσκεται υπό πίεση" übersetzt werden kann.



Lexikalische Lücken-1

- Typische Beispiele von Nichtenstprechungen oder "Lexikalischen Lücken,, bezüglich des Sprachpaars Deutsch-Griechisch sind die Wendungen "Muster", "Gewalt,, und "Zeit".
- Die Wendung "**Muster**" kann im Griechischen als "σχέδιο", als "δείγμα" sowie auch als "πρότυπο,, interpretiert werden.
- Das Wort "**Gewalt**" kann im Griechischen "εξουσία" (Macht), "δύναμη" (Kraft), "έλεγχος" (Kontrolle) oder auch "βία" (Gewalttätigkeit) bedeuten.



Lexikalische Lücken-2

- Im Falle einer Übersetzung vom Griechischen in die deutsche Sprache muss daher das zur Unterscheidung notwendige Wissen (Dorna und Jekat, 2004) aus dem Kontext erschlossen werden.
- Die Wendung "Zeit" ist ein weiteres Beispiel lexikalischer Lücken. Es handelt sich um ein polysemes Wort, das im Griechischen als "καιρός, χρόνος", als "καιρός, ώρα", als "εποχή" oder als "χρόνος" interpretiert werden kann.
- Die Probleme, die aus den Nichtentsprechungen entstehen können, werden komplexer, wenn beide Wendungen in der gleichen Subsprache, d.h. im gleichen Kontext erscheinen können.



Subsprachen und kontrollierte Sprachen

- Misslungene Übersetzungen können zu einem relativ hohen Grad vermieden werden, wenn das Anwendungsgebiet der maschinellen Übersetzung eingegrenzt wird.
- Je beschränkter (und spezialisierter) die Domäne bzw. das Anwendungsgebiet des Systems sprachlicher Verarbeitung ist, desto weniger und einfachere Regeln hat es, und desto geringer sind die Chancen, misslungene Übersetzungen zu erzeugen.
- Die spezialisierte und/oder beschränkte Sprache der Domäne wird auch **Subsprache** genannt.



Kontrollierte Sprachen

- Um die Domäne auch linguistisch zu beschränken, werden **kontrollierte Sprachen** benutzt.
- Kontrollierte Sprachen werden durch eine Menge lexikalischer, semantischer, morphosyntaktischer und pragmatischer Regeln bestimmt.
- Die ersten kontrollierten Sprachen wurden für die Verfassung und Übersetzung von Texten der Luft- und Raumfahrtindustrie (Boeing und Aerospatiale) geschaffen.
- Kontrolliertes Deutsch wird z.B. in der technischen Dokumentation benutzt (Lehrndorfer, 1996).



Typische lexikalische und syntaktische Beschränkungen-1

für deutsche **kontrollierte Sprachen**:

Typische Beispiele für allgemeine Richtlinien (nach Lehrndorfer, 1996):

1. WORTSCHATZ

1.1. TERMINOLOGIE

Nur Terminologie des definierten Fachbereichs verwenden

1.2. ALLGEMEINER

WORTSCHATZ

Wortschatz der Sprache des definierten Fachbereichs verwenden
Polysemie bezüglich der Wörter und Wendungen vermeiden



Typische lexikalische und syntaktische Beschränkungen-2

2. SYNTAX

2.1. SATZSTRUKTUR Kurze Sätze konstruieren

Elliptische Sätze und elliptische Phrasen verwenden

2.2. GRAMMATISCHE KATEGORIEN

Polysemie bezüglich der Negationen vermeiden

Partizipialkonstruktionen vermeiden

Ambiguitäten bezüglich des Zustands und des Vorgangspassivs vermeiden

Imperativischen Infinitiv verwenden



Was sind Dialogsysteme?

- Natürlichsprachliche Dialogsysteme erlauben es einem menschlichen Benutzer mit einer Maschine mittels sprachlicher Ein-und Ausgabe zu kommunizieren (Kellner, 2004).
- Üblicherweise ist mit der Wendung "Dialogsystem,, ein System zur Verarbeitung gesprochener Sprache gemeint, mit dem der Benutzer mündlich kommuniziert und von dem System Antworten oder Reaktionen in der Form gesprochener Sprache bekommt.



Grundkomponenten eines Dialogsystems

- Die meisten Dialogsysteme enthalten drei Grundkomponenten für die Verarbeitung der gesprochenen Sprache:
 - (1) die Spracherkennungs und Sprachverstehenskomponente,
 - (2) die Dialogsteuerung und
 - (3) das Sprachgenerierungsmodul.



Spracherkennung und –verstehen (1)

- Über die Spracherkennungs- und Sprachverstehenskomponente findet die **Analyse** der eingehenden Benutzeräußerungen statt.
- Die Benutzeräußerungen werden anschließend in eine **semantische Darstellung** überführt (Kellner, 2004).
- Die Spracherkennungs- und Sprachverstehenskomponente soll die Stimmen und Aussprachen von mehr als einem Benutzer erkennen, also **sprecherunabhängig** sein ("user-independent"), und zwar ohne, dass der Benutzer das Dialogsystem in Bezug auf seine Stimme trainiert hat.



Dialogsteuerung (2)

Dialogsteuerung und pragmatische Regeln

- Bei komplizierten und/oder multimodalen Systemen sprachlicher Verarbeitung wie den Dialogsystemen werden auch pragmatische Regeln benutzt.
- Programme die pragmatische Regeln aktivieren, gehören zu den schwierigsten und anspruchsvollsten Anwendungsgebieten der Informatik und der künstlichen Intelligenz.



Die pragmatischen Regeln kontrollieren die Sprechakte,

- sind meistens von den Grammatiken unabhängig und haben die Form von Algorithmen.
- Diese arbeiten mit Programmen die den ganzen Text bzw. Dialog behandeln (Dialog Manager, Shriberg et al., 2002 (NASA-SRI)).
- Diese Programme spielen die Rolle des "Schiedsrichters" oder "Verkehrspolizisten,, in dem vom System zu verarbeitenden Text.
- Dieser Prozess wird auch "Dialogsteuerung" ("dialog-management") genannt.
- Die Dialogsteuerung gibt die Systemäußerungen nur in abstrakter semantischer Form vor. Die Umsetzung der von der Dialogsteuerung in abstrakter semantischer Form vorgegebener Systemäußerungen findet in dem Sprachgenerierungsmodul statt.



Sprachgenerierung und –synthese (3)

- Das Sprachgenerierungsmodul setzt die von der Dialogsteuerung in **abstrakter semantischer Form** vorgegebenen Systemäußerungen in eine textuelle Repräsentation um.
- Die textuelle Repräsentation der Systemäußerungen wird dann mit Hilfe eines **Sprachsynthesystems** in gesprochene Sprache umgesetzt und an den Benutzer ausgegeben.
- Oft wird die Qualität der Ausgabe des Systems durch zusätzliche Mittel verbessert, zum Beispiel in Form von prosodischer Modellierung (wie die Verwendung von Prosodiemarkern).



Beispiel (Alexandris, 2010)

BENUTZER: Wie ist das Wetter in Bayern?

COMPUTER (SYSTEM): Im Alpengebiet Schnee, Temperatur minus zwei Grad, München bewölkt, leichter Regen, Temperatur fünf Grad, Nürnberg heiter bis wolkig, Temperatur sieben Grad, leichter Wind.

BENUTZER: Wie ist das Wetter in München?

COMPUTER (SYSTEM): Das Wetter heute in München ist bewölkt, leichter Regen, Temperatur fünf Grad, abends zwei Grad.

BENUTZER: Wie ist das Wetter in Garmisch-Partenkirchen?

COMPUTER (SYSTEM): Das Wetter heute im Alpengebiet Schnee, im Bereich der Zugspitze Schneeschauer, Temperatur minus zwei Grad, abends minus fünf Grad, im Bereich der Zugspitze, leichter Wind.

BENUTZER: Wie wird das Wetter morgen in München sein?

COMPUTER (SYSTEM): Das Wetter für morgen in München ist heiter bis wolkig, Temperatur fünf bis sieben Grad, leichter Wind.



Beispiel (Alexandris, 2010)-2

- (a)
- GEBIET: Bayern
Ort: München:
RELATIVE-ZEIT: heute
TAGESZEIT: (jetzt):
WETTER:
Himmel: bewölkt
Niederschläge: leichter Regen
Temperatur: fünf Grad
TAGESZEIT: abends
Temperatur: zwei Grad
RELATIVE-ZEIT: morgen
TAGESZEIT (unbestimmt):
WETTER:
Himmel: heiter bis wolkig
Temperatur: fünf bis sieben Grad
Wind: leichter Wind
- (b)
- Wetterbericht für: München,
Bayern
- Heute: bewölkt, leichter Regen
Temperatur: fünf Grad
abends: Temperatur: zwei Grad
morgen: heiter bis wolkig,
leichter Wind
Temperatur: fünf bis
sieben Grad



Greece: CitizenShield Project

www.polias.gr (Nottas et al, 2007)

Beispiel (Dialogsystem-Griechisch)

[7]: INTERACTION 7: SYSTEM: Το προϊόν το αγοράσατε από **σούπερμάρκετ**, από **μικρό κατάστημα**, από **λαϊκή αγορά**, από **περίπτερο** ή από **κυλικείο**;

BENUTZER: «ΚΑΤΑΣΤΗΜΑ»

«ΚΑΤΑΣΤΗΜΑ» [8.1] (* για άλλες επιλογές)

[8.1]: INTERACTION 8: SYSTEM: Το είδος σας **εστάλη** από το κατάστημα;

BENUTZER: «ΟΧΙ»

[8.2]: INTERACTION 8: SYSTEM: Πώς **ονομάζεται** το κατάστημα και σε **ποια** περιοχή έγινε η αγορά;

BENUTZER: NAME-ORT

[8.3]: INTERACTION 8: SYSTEM: Πείτε μας, αν ξέρετε, τη **διεύθυνση** του σημείου αγοράς;

BENUTZER: ADRESSE

[9]: INTERACTION 9: SYSTEM: Πείτε μα αν θυμάστε, ποια **ημερομηνία** έγινε η αγορά;

BENUTZER: DATUM



Literaturverzeichnis

- Alexandris, C. (2010): Linguistik und ihre Anwendungen in der Computerlinguistik: Ein Arbeitsbuch, Athens, Papatotiriou. (Students book, in German)
- Alexandris, C., Fotinea, S-E and Efthimiou, E. (2005). Emphasis as an Extra-Linguistic Marker for Resolving Spatial and Temporal Ambiguities in Machine Translation for a Speech-to-Speech System involving Greek. In: *Proceedings of the 3rd International Conference on Universal Access in Human-Computer Interaction (UAHCI 2005)*, 22-27 July 2005, Las Vegas, Nevada, USA.
- Alexandris, C. (2003). Translational Issues in the Sublanguage of Written and Spoken Journalistic Texts in Modern Greek. In: *Proceedings of the International Conference on Choice and Difference in Translation*, Athens 2003, 287-307.
- Bateman, J, Paris, C. (2004). Benutzermodellierung. In: *Computerlinguistik und Sprachtechnologie, Eine Einführung*, Carstensen, K.U., Ebert, C., Endriss, C., Jekat, S., Klabunde, R., Langer, H. (Hrsg.), 2te überarbeitete und erweiterte Auflage, München: Spektrum Akademischer Verlag.
- Cohen, P., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L., and Clow, J. (1997). Quickset: Multimodal interaction for distributed applications. In: *Proceedings of the 5th ACM International Multimedia Conference*, pages 31-40.
- Dorna, M., Jekat, S. (2004). Maschinelle und computergestützte Übersetzung. In: *Computerlinguistik und Sprachtechnologie, Eine Einführung*, Carstensen, K.U., Ebert, C., Endriss, C., Jekat, S., Klabunde, R., Langer, H. (Hrsg.), 2te überarbeitete und erweiterte Auflage, München: Spektrum Akademischer Verlag.
- Forrester, M. (1996). *Psychology of Language*. SAGE Publications, Thousand Oaks, CA, USA.
- Hatim, B. (1997). *Communication Across Cultures: Translation Theory and Contrastive Text Linguistics*, University of Exeter Press.
- Jurafsky, D., Martin, J. (2008). *Speech and Language Processing, an Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*, 2nd edition, Prentice Hall series in Artificial Intelligence, Pearson Education, Upper Saddle River, NJ, USA.



Literaturverzeichnis

- Kellner, A. (2004). Dialogsysteme. In: *Computerlinguistik und Sprachtechnologie, Eine Einführung*, Carstensen, K.U., Ebert, C., Endriss, C., Jekat, S., Klabunde, R., Langer, H. (Hrsg.), 2te überarbeitete und erweiterte Auflage, München: Spektrum Akademischer Verlag.
- Lehrndorfer A. (1996). *Kontrolliertes Deutsch: Linguistische und Sprachpsychologische Leitlinien für eine (maschniell) kontrollierte Sprache in der technischen Dokumentation*, Tübingen: Narr.
- Moegele, H., Moritz Kaiser, M., Schiely, F. (2006). SmartWeb UMTS Speech Data Collection, The SmartWeb Handheld Corpus. In: *Proceedings of LREC 2006*, Genova, Italy, pp. 2106-2111.
- Müller, S. (1998): Babel 1.50, Web-Interface, Universität Bremen.
- v. Hahn, W. (2001). Maschinelle Übersetzung, Proseminar der Fakultät für Informatik, Universität Hamburg.
- Hanneforth, T. (2001). Was ist Computerlinguistik?, Übersicht des Computerlinguistikprogramms, Institut für Linguistik, Universität Potsdam.
- Shriberg, E, Stolcke, A., Stone, L., Bratt, H., Ferrer, L. and Sömnez, K. (2003). Harnessing Speech Prosody for Robust Human-Computer Interaction, Active Research Task, Intelligent Systems Project, CICT, SRI-International, NASA-Ames Research Center.
- Tomita, M., Mitamura, T., Musha, H. and Kee, M. (1988). The Generalized LR Parser/Compiler Version 8.1, Center For Machine Translation, Carnegie Mellon University, Pittsburgh, PA, USA
- Wardhaugh, R. (1992). *Introduction to Sociolinguistics*. Oxford, Blackwell.
- Wodack, R (1996). *Disorders of Discourse*. Longman, New York.
- Vertan, C. (2001). Einführung in Grundprobleme der Maschinellen Übersetzung, Seminar der Fakultät für Informatik, Universität Hamburg.
- Έργο διαλογικού συστήματος Verbmobil (Γερμανία) <http://verbmobil.dfki.de/>



Τέλος Ενότητας

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Σημειώματα

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.0.

Έχουν προηγηθεί οι κάτωθι εκδόσεις:

- Έκδοση διαθέσιμη εδώ. <http://eclass.uoa.gr/courses/GS158/>



Σημείωμα Αναφοράς

Copyright Εθνικών και Καποδιστριακών Πανεπιστημίων Αθηνών, Χριστίνα Αλεξανδρή. «Υπολογιστική Γλωσσολογία. Multilinguale Mensch-Maschine Kommunikation: Linguistische Aspekte und Anwendungen-1». Έκδοση: 1.0. Αθήνα 2014. Διαθέσιμο από τη δικτυακή διεύθυνση:
<http://opencourses.uoa.gr>



Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.



Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.



Σημείωμα Χρήσης Έργων Τρίτων (1/2)

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Εικόνες/Σχήματα/Διαγράμματα/Φωτογραφίες



Σημείωμα Χρήσης Έργων Τρίτων (2/2)

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Πίνακες

