

Architekturen von KI-Systemen

Architektur des Verbmobil-Systems

Muhammet Ali Salur

Katarzyna Bade

David Ortiz Luque

Proseminar SS04 Cristina Vertan

Inhalt (I)

- Was ist Verbmobil?
- Der Forschungsprototyp von Verbmobil
- Systemintegration
- Sprachsynthese
- Spracherkennung
 - Einführung
 - Mehrsprachige Spracherkennung
 - Spracherkennung in Verbmobil
 - Identifizierung der Sprache
- Robuste Erkennung der Spontane Sprache
 - Die Herausforderung des Geräusches
 - Die Herausforderung der Transmission-Kanälen
 - Die Herausforderung des Sprechers

Inhalt (II)

- Schnelle Suche für große Wortschatz Spracherkennung
- Der „prosodic“ Modul
- Warum ist Prosodie benutzen schwierig?
- Transfer und Übersetzung
 - Transferkonzepte für Vermobil
 - Die Dolmetschfunktion für Vermobil
 - Bestimmung von Transfereinheiten eines Dialoges
- Beispiele
- Das Handlungschema – Terminvereinbarung
- Transfer - eine Kontrollinstanz in Vermobil
- Literaturverzeichnis

Was ist Verbmobil?

„Verbmobil“ ist der Name eines Projekts, in dem ein System entwickelt wird, das natürliche, spontan gesprochene Sprache simultan dolmetschen soll.

Mit „Dolmetschen“ ist dabei nicht die wörtliche Übersetzung der Sätze gemeint.

Das Verbmobil-System:

erkennt gesprochene Spontansprache,

analysiert die Eingabe,

übersetzt sie in eine Fremdsprache,

erzeugt einen Satz,

spricht ihn aus.

Für ausgewählte Themenbereiche, z.B. Terminverhandlung, Reiseplanung, Fernwartung, soll Verbmobil Übersetzungshilfe in Gesprächssituationen mit ausländischen Partnern leisten.

Die Komplexität des Themenbereichs erfordert eine interdisziplinäre Zusammenarbeit.

Die Beteiligten kommen aus den Disziplinen Phonetik, Syntax, Semantik, Logik, Übersetzungswissenschaft, Nachrichtentechnik und Soziologie. Hinzu kommt noch die Ebene der Einzelphilologien: Anglistik, Germanistik, Japanologie.

Unter der Leitung des Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI GmbH), arbeiten an diesem Verbundprojekt zahlreiche Universitäten, Forschungszentren und Unternehmen zusammen.

Industriepartner sind: Daimler-Benz, DASA, Philips, Siemens.

Forschungspartner sind: DFKI, LMU München, RWTH Aachen, TU Berlin, TU Dresden, TU München, Univ. Bielefeld, Univ. Bochum, Univ. Bonn, Univ. Braun-schwieg, Univ. d. Saarlandes, Univ. Erlangen, Univ. Hamburg, Univ. Karlsruhe, Univ. Stuttgart, Univ. Tübingen

Vom Fachbereich Informatik der Univ. Hamburg ist der Arbeitsbereich NATS beteiligt.

Gefördert wird das Verbmobil vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF).

Nach der Laufzeit wird das Projekt in zwei Phasen eingeteilt.

Geplant war die erste Phase von 1993 bis 1996 und finanziert vom BMBF mit 64,9 Mio. DM und aus der Wirtschaft mit 31 Mio. DM.

Die Ziele der ersten Stufe wurden erreicht, und so wurde die zweite Phase (1997-2000) gestartet, gefördert mit 50,2 Mio. DM vom BMBF und 20,4 Mio. aus der Wirtschaft.

Der Forschungsprototyp von Verbmobil

Alle technischen Ziele der ersten Phase von Verbmobil wurden voll erreicht und in einem Forschungsprototypen realisiert:

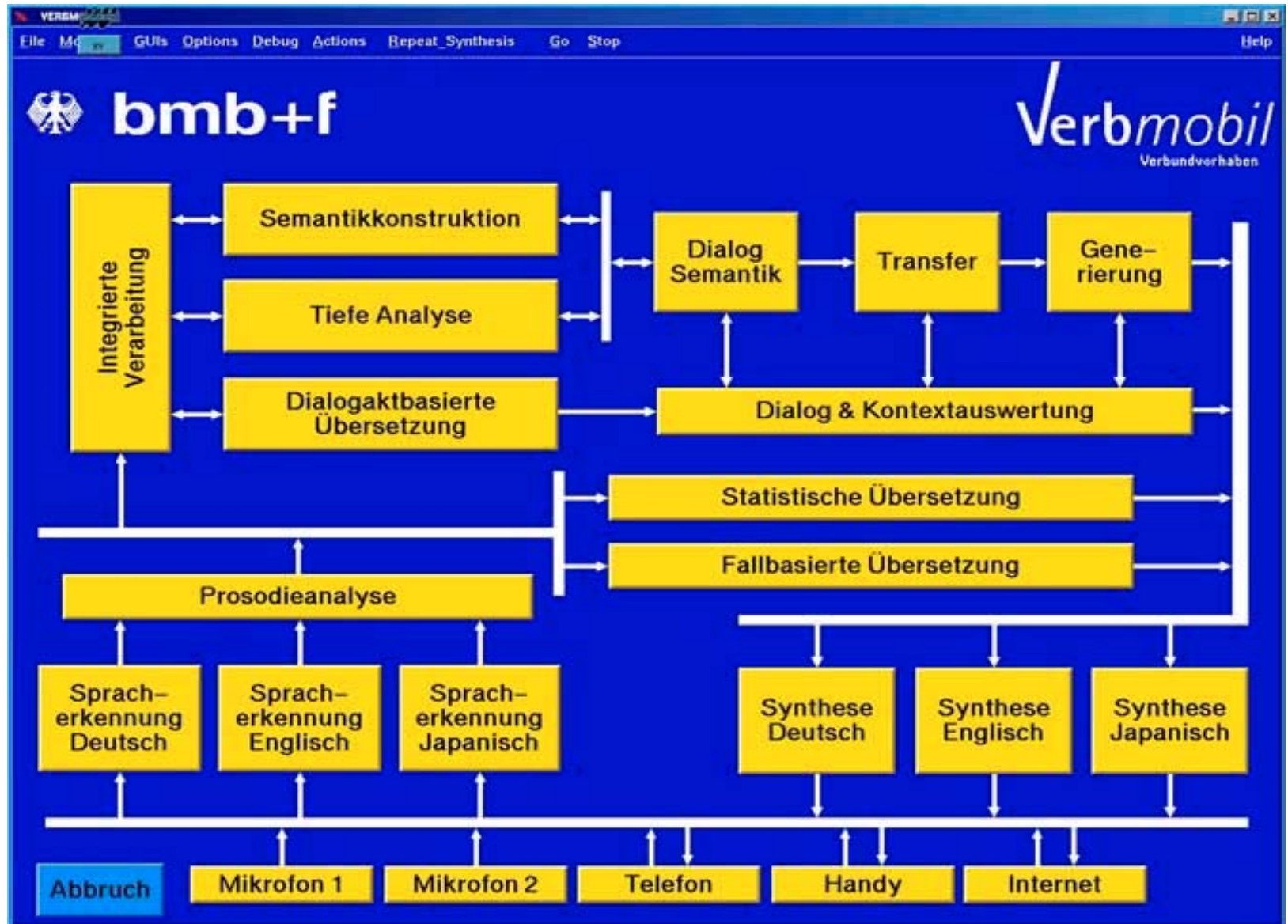
1. Erkennung fließend gesprochener Spontansprache für Deutsch, Japanisch und Englisch über Nahbesprechungsmikrofon
2. Wortschatz von ca. 2500 Wörtern für die Übersetzungsrichtung Deutsch nach Englisch
3. Sprecheradaptives System mit sprecherunabhängigem Kern
4. Linguistisch fundierte deutsche Basisgrammatik für Spontansprache mit tiefer und flacher semantischer Analyse
5. Gesprochene Klärungsdialoge zwischen dem Benutzer und dem Verbmobil- System bei Spracherkennungs- und Verstehensproblemen

6. Semantischer Transfer für Deutsch -Englisch und Japanisch -
Englisch
7. Sprachgenerierung für Englisch und für deutsche Paraphrasen
8. mehr als 70% approximativ korrekte Übersetzungen bei der
End-to-End Evaluation in der Domäne Terminverhandlung
9. Reine Softwarelösung für alle Module auf Standardhardware
10. Netto-Verarbeitungszeit < sechsfache Echtzeit bezogen auf die
Länge des Eingabe-Sprachsignals

Die zweite Phase von Verbmobil kann zusammenfassend wie folgt gekennzeichnet werden:

1. **Multifunktionalität:** Verbmobil soll rasch auf neue Gesprächsdomänen einstellbar sein.
2. **Multilingualität:** Verbmobil soll spontane Dialoge in mehrere Sprachen übersetzen können.
3. **Multimedialität:** Verbmobil soll in internationalen Multimedia-Anwendungen Übersetzungshilfe anbieten.
4. **Mobilität:** Durch einen Sprachserver soll Verbmobil auch über Handy nutzbar sein.
5. **Multiparty-Funktionalität:** Verbmobil soll nicht nur in Dialogsituationen, sondern auch in Telekooperationsanwendungen mit vielen Gesprächspartnern Übersetzungshilfe leisten.

Fig. 1: Die Benutzeroberfläche des Forschungsprototypen von Verbmobil



Systemintegration

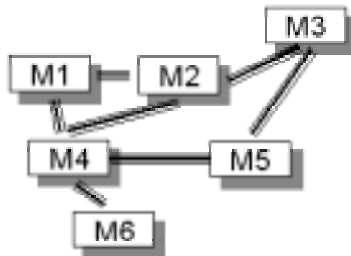
Wie die Architekturübersicht in Fig. 1 zeigt, wurde Verbmobil als hochgradig nebenläufiges System nach dem **Multiagenten-Prinzip** mit zahlreichen Kommunikationsschnittstellen zwischen den Verarbeitungsmodulen vollständig objektorientiert realisiert.

Die Benutzeroberfläche, durch die auch der Verarbeitungsablauf visualisiert wird, zeigt nur die Hauptmodule der insgesamt 43 Systemkomponenten.

In Verbmobil kommunizieren Module, die in Fortran, C, C++, Lisp und Prolog implementiert wurden, in einem objektorientierten Architekturkonzept für Software. Eine zentrale Gruppe für die Verbmobil Systemintegration am Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI GmbH) überprüft die Module und integriert sie zu einem Gesamtsystem.

Verbmobil I

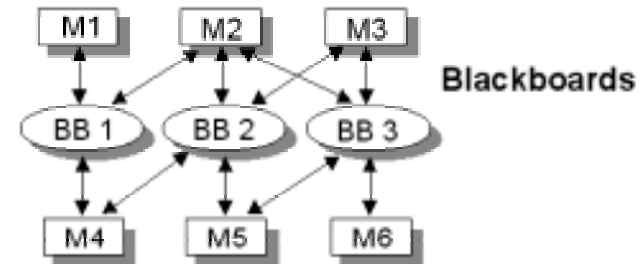
- Multi-Agent Architecture



- Each module must know, which module produces what data
- Direct communication between modules
- Each module has only one instance
- Heavy data traffic for moving copies around
- Multiparty and telecooperation applications are impossible
- Software: ICE and ICE Master
- Basic Platform: PVM

Verbmobil II

- Multi-Blackboard Architecture



- All modules can register for each blackboard dynamically
- No direct communication between modules
- Each module can have several instances
- No copies of representation structures (word lattice, VIT chart)
- Multiparty and Telecooperation applications are possible
- Software: PCA and Module Manager
- Basic Platform: PVM

Fig. 2: Ein Vergleich der Architekturen von Verbmobil I und II

Sprachsynthese

Sprachsynthese, genauer: Stimmsynthese, hat die Aufgabe, eine künstlich sprechende Stimme zu erzeugen.

Es gibt drei grundsätzlich verschiedene Richtungen in der modernen Sprachsyntheseforschung.

Sie unterscheiden sich in ihren Zielen voneinander.

- TEXT-TO-SPEECH (TTS)
- ARTIKULATORISCHE SPRACHSYNTHESE
- CONCEPT-TO-SPEECH (CTS)

Text-to-Speech (TTS) bezeichnet die Umwandlung von schriftlich vorhandenem Text in sprachliche Äußerungen. Sehr wichtig hierbei ist die linguistische Vorverarbeitung des Materials.

Die **artikulatorische Synthese** wird in der heutigen Phonetik zur Grundlagenforschung verwendet und findet keinen Einsatz im kommerziellen Bereich.

Concept-to-Speech (CTS, auch Speech Generation) ist die Erzeugung von Information aus Konzepten oder abstrakten Daten und die entsprechende Generierung der lautlichen Äußerung.

Deswegen sind hier auch ganz andere Gesichtspunkte entscheidend:

Fragen der Art „Was will der Hörer und welches Vorwissen hat er?“ sind maßgeblich für systeminterne Entscheidungen.

Welche Art der Sprache muß gewählt werden? Welches sind die bedeutungstragenden Elemente innerhalb der Information?

Gerade bei CTS sind Prosodie und die richtige Akzentuierung von immenser Bedeutung.

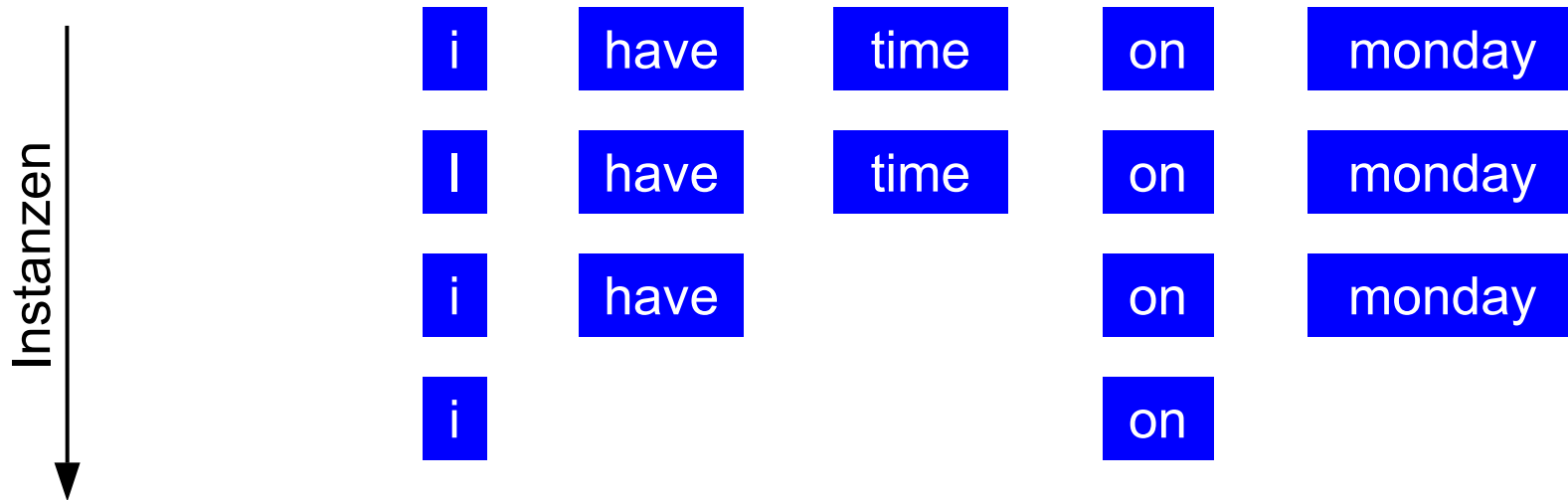
Es wird zwischen zwei unterschiedlichen Arten der Synthese unterschieden:

- **regelbasiert**
- **korpusbasiert**

Verbmobil besitzt eine korpusbasierte CTS-Synthese

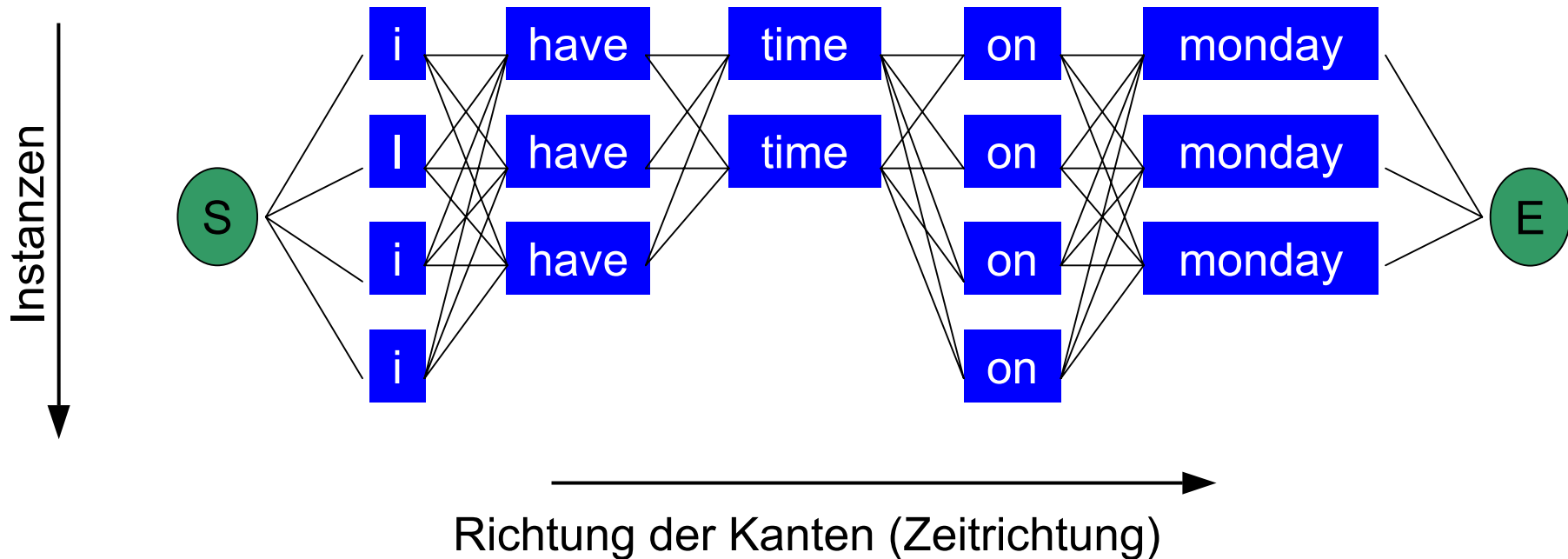
- Bei dieser Methode werden ganze Bausteine (sog. units) einer Sprache aufgenommen und dann anhand der Symbolketten modifiziert und neu zusammengesetzt.
- Das **Wort ist zentrale Einheit** und Ausgangspunkt für jede Verarbeitung.
- Es werden keine isolierten Sprachlaute aufgenommen, sondern Lautübergänge, also z.B. Diphone.

Zieläusserung: **I** **have** **time** **on** **Monday.**



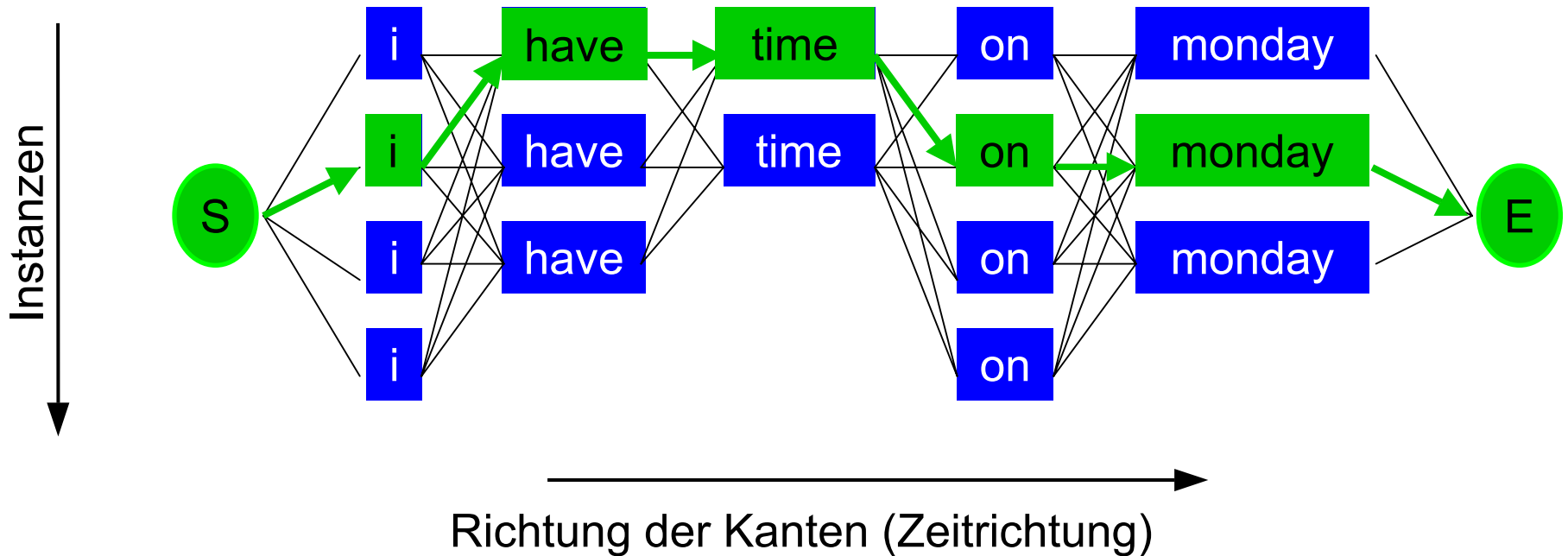
In einem ersten Schritt werden alle (passenden) Instanzen der Wörter der Zieläusserung in einer Matrix zusammengestellt.

Zieläusserung: **I have time on Monday.**



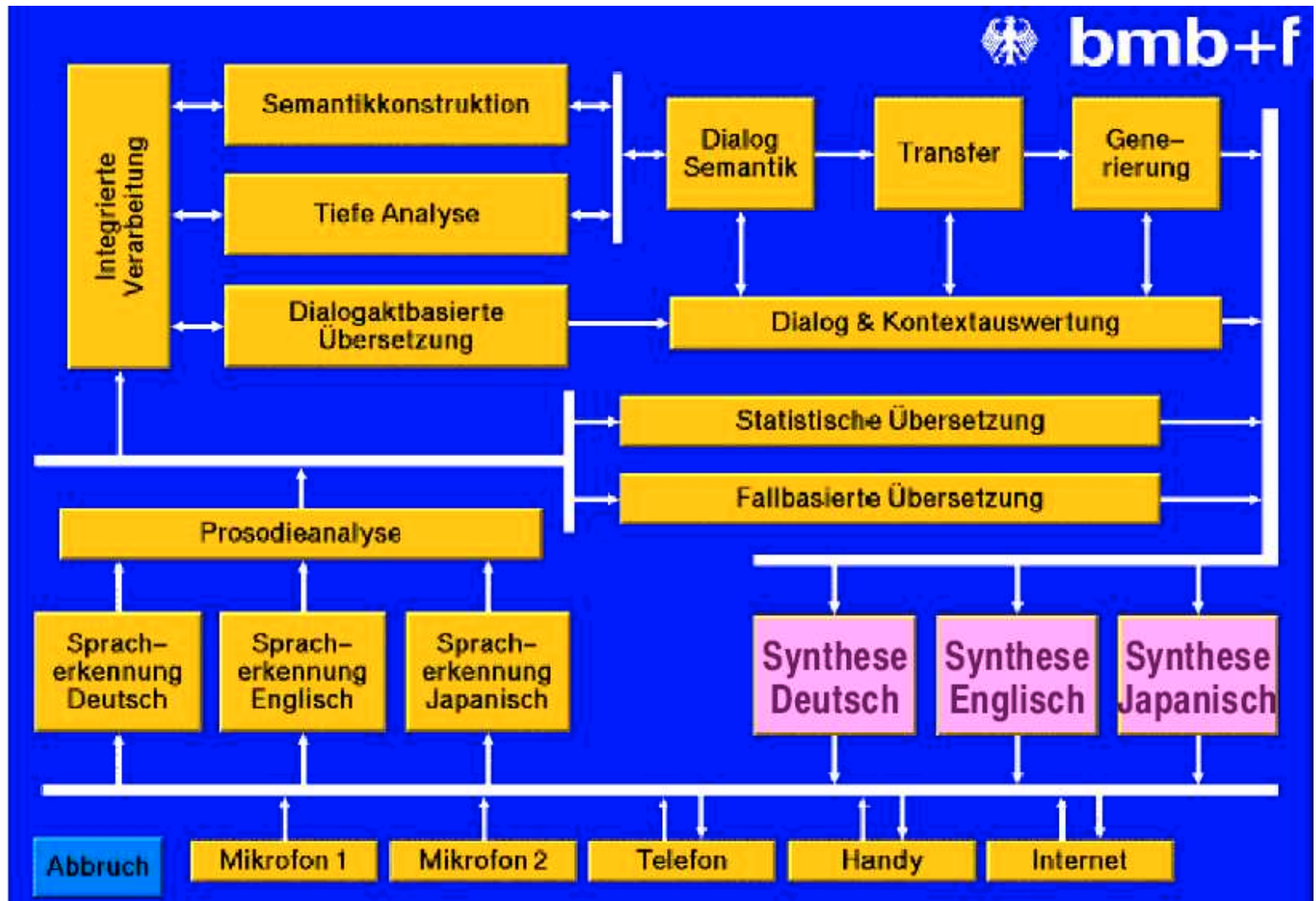
Alle Instanzen in der Matrix werden zu einem gerichteten Graphen zusammengefasst; dieser wird mit einem Start- und Endknoten versehen.

Zieläusserung: **I** **have** **time** **on** **Monday.**



Die Auswahl der tatsächlich benutzten Bausteine erfolgt durch Ermittlung eines optimalen Pfades durch den Graphen.

Anbindung der Synthese in Verbmobil



Ein CTS-System bezieht die Information über die höheren Ebenen (Syntax, Semantik) aus den zentralen Modulen des übergeordneten Systems, von dem es ein Teil ist.

Es ist zu erwarten, dass durch die zusätzliche semantische und pragmatische Information Aussprachefehler vermieden werden und die Gesamtqualität insbesondere im Bereich der Prosodie (Satzbetonung, Lautdauern, Pausen) steigt.

Problem ist das Durchreichen der Information.

Die herkömmlichen symbolverarbeitenden Komponenten werfen Information weg, wenn sie nicht mehr benötigt wird, und nehmen keine Rücksicht auf nachfolgende Module.

Für den schlimmsten Fall muss TTS als Rückfallposition bereit stehen, falls die Zusatzinformation ausfällt.

Spracherkennung

Spracherkennung: Einführung (I)

- Verbmobil ist bis heute das einzigste System der „Dialog Übersetzung“.
- Basierend auf einer „offenen Mikrofon“ Bedingung.
- Es ist kein „push-to-talk“ System.
- Das Signal kann aus verschiedene Qualitäten bestehen (Mikrofon Geschlossen-Sprechen mit Lab-Qualität, GSM Qualität, ...).
- Verbmobil arbeitet mit spontaner Sprache.
- Im Verbmobilcorpus: 20% Dialoge haben mindestens eine Selbst-Korrektur und 3% Dialoge beginnen falsch.
- Verbmobil benutzt viele flache und tiefe Analysis-Methoden um einen Sprecherlapsus zu erkennen und kann es im voraus übersetzen.

Spracherkennung: Einführung (II)

- Am Anfang benutzt man „prosodic“ Zeichen um Selbst-Korrekturen zu erkennen.
- Ein stochastisch Modell trennt den falschen Teil (das heißt reparandum) und die Korrektur voneinander. Dann ist die korrigierte Eingabe wie eine neue Hypothese eingefügt.
- Die Reparierte-Verarbeitung in Verbmobil ist ein Filter zwischen Spracherkennung und syntaktische Analyse.
- Es gibt ein Modul, das mit den Teilen der Worte arbeitet, die der Sprecher zu korrigieren versucht und beseitigt Sounds wie „ahh“ (Benutzer machen das oft während sie sprechen).
- Verbmobil ist ein Sprecher-anpassendes System.
- Die Aufnahme des Verbmobilcorpus: In Parallele mit Mikrofon Geschlossen-Sprechen, Raum-Mikrofon und einige Telefone (GSM, drahtlos, DECT und regulär).

Mehrsprachige Spracherkennung

- Verbmobil ist ein mehrsprachiges Übersetzungssystem „Sprache-Sprache“.
- Drei Sprachen (Deutsch, Englisch und Japanisch): 700 Millionen gebürtige Sprecher in der Welt.
- Das System erkennt automatisch die Eingabesprache des Benutzers und benutzt den entsprechenden Motor.
- Unterschiede in der Sprache:
 - Buchstaben und Fonts
 - Romanization und Segmentierung
 - Prosodish Struktur
 - Morphologie

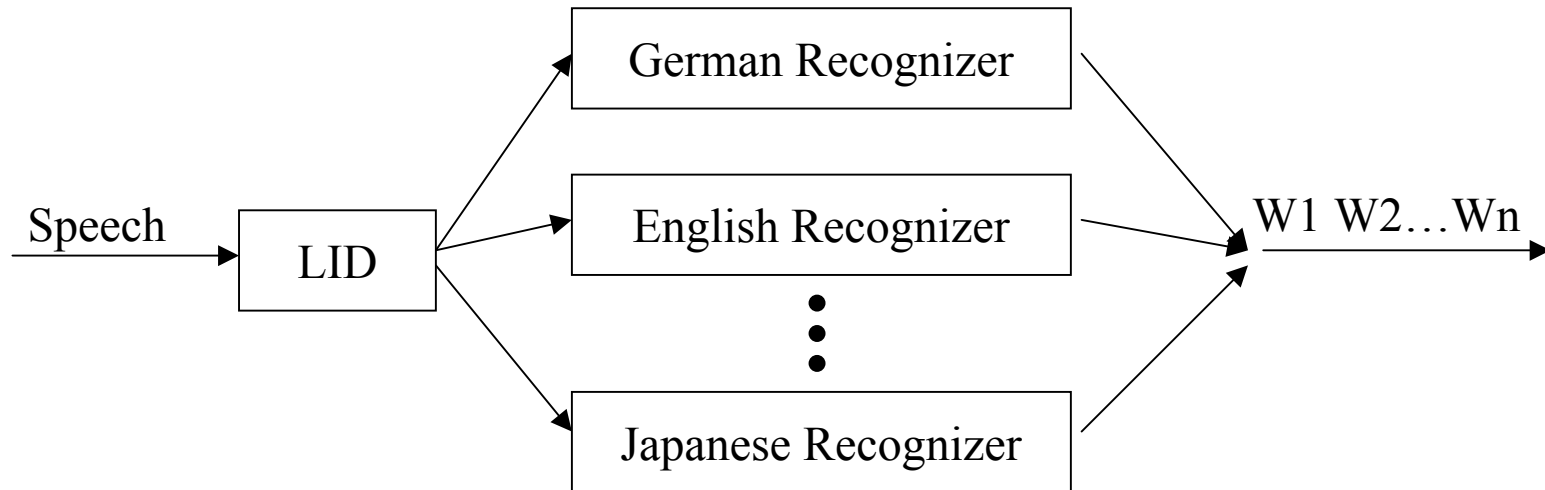
Spracherkennung in Verbmobil

Item	German	English	Japanese
Training Data	62 hrs	32 hrs	39 hrs
Vocabulary (word forms)	10254	7965	3490
Pronunciation variants per word	1.15	1.20	1.08
Phenomene inventory	47	41	32
Acoustic Model (Quinphones)	3300	2250	2500
Language Model Corpus	670K	270K	580K
Trigram Testset Perplexity	77	47.3	17.3
Testset OOV-rate	1.0%	1.0%	2.6%

Identifizierung der Sprache (I)

- Es gibt drei Motoren (einen für jede Sprache) um die Sprache des Benutzers direkt zu erkennen. So das, dass System flexibel und Benutzerfreundlich ist.
- Der Bestandteil „Language IDentification“ (LID): Es gibt zwei verschiedene Annäherungen an dieses Problem:
 - Score-based: Man benutzt eine mehrsprachige Erkennung oder verschiedene einsprachige Erkennungen und man rechnet die Punktzahl (eine Nummer zeigt „wie passend die akustischen und gespeicherten Models“ je Sprache sind). Die bessere Punktzahl bestimmt die Sprache.
 - Confidence-based: Man teilt jedem Erkennen einen Sicherheitswert zu. Die mit der höchsten Sicherheit bestimmt die Sprache.

Identifizierung der Sprache (II)



Robuste Erkennung der Spontane Sprache

- Die Herausforderung des Geräusches
 - Die Erhöhung der Sprache und die Entschädigung des Geräusches: Algorithmen SPS und PMC
 - Optimierung und Implementierung der Robusten Methoden in Verbmobil
- Die Herausforderung der Transmission-Kanälen
 - Transmission Kanälen in Verbmobil
- Die Herausforderung des Sprechers
 - Normalisierung Methoden
 - Akustische Anpassung
 - Lexikalische Anpassung
 - Sprechen Steuer

Die Herausforderung des Geräusches (I)

- Verbmobil fordert robuste Spracherkennung in verschiedenen Umgebungen (ruhiger Raum ohne Geräusche oder via Handy in unbekannter Umgebung oder an einem Platz mit vielen Leuten und vielen Geräuschen).
- Man versucht nicht Beispiele für alle möglichen geräuschvollen Situationen zu finden um das System zu verbessern.
- Man trainiert das System mit Geräuschfreien Beispielen und man wendet Methoden an um den Effekt der Geräusche zu verringern.
- Das System funktioniert in Situationen aller Art (mit / ohne Geräusch).
- Die Technik werden in drei Kategorien eingestuft:
 - Verwendung von Geräusch robusten Eigenschaft
 - Erhöhung der geräuschvollen Sprache
 - Entschädigung der Sprachemodelle in geräuschvollen Umgebungen.

Die Herausforderung des Geräusches (II)

- Die Erhöhung der Sprache und die Entschädigung des Geräusches:
 - Algorithmus SPS (Spectral Substraction): Man schätzt das Geräusch in Sprachpausen und man entfernt es von dem Eingangssignal.
 - Algorithmus PMC (Parallel Model Combination): Der PMC säubern das Sprachsignal nicht, sondern er gliedert das Geräusch in der Erkennung ein.
- Optimierung und Implementierung der Robusten Methoden in Verbmobil:
 - Das Schema der Modellkombination benutzt die Technik „Linear Discriminant Analysis“ (LDA) and „Cepstral Mean Substracion“ (CMS).
 - Die Eigenschaften des Geräusches berechnet man nicht im Voraus, sondern sie schätzt man online während die Erkennung ein und man aktualisiert fortlaufend. Das ist wichtig in unbekanntem oder geänderten Umgebungen.

Die Herausforderung des Geräusches (III)

- Man braucht robusten Methoden um befriedigenden Ergebnisse in unbekannten oder geänderten Umgebungen zu erreichen (nicht nur in ruhigen Umgebungen).
- Die Experimente zeigen bedeutsamen Besserungen für diese Methoden (wenn man sie in geräuschvollen Umgebungen benutzt).
- Aber man kann nicht so gute Erkennung wie in ruhigen Umgebungen bekommen.

Die Herausforderung der Transmission-Kanälen

- Verbmobil benutzt man in verschiedenen Umgebungen mit verschiedenen Eingangsvorrichtung. Jede Vorrichtung beeinflusst das Sprachsignal irgendwie.
- In Verbmobil hat man drei Kanälen um das System zu trainieren und um das zu testen benutzt: Ein Kopfhörer Mikrofon (Hoche-Qualität Aufnahmen), ein Desktop Mikrofon und die Telefonleitung (Analog, drahtlos und Handy).
- In Verbmobil ist das Handy der Wichtigste (man kann das System irgendwo erreichen).
- Man hat zwei verschiedene Erkennen um das Verbmobil System zu auswerten trainiert: Mit dem Kopfhörer Mikrofon und mit der Telefonleitung.
- Man hat diese zwei Erkennen separat errichtet und optimiert.

Die Herausforderung des Sprechers (I)

- Es gibt Situationen, wo Menschen Probleme um ihre Partner in Anfangsgespräch zu verstehen gehabt haben (der Gegensprecher spricht zu schnell, zu laut, zu ruhig, zu nachlässig, Umgangssprache oder einen Dialekt).
- Aber nach einige Sekunden versteht man erfolgreicher. Dieses Verfahren heißt „Adaptation“ (Anpassung) und man muss in einem nicht-menschlichen Erkennen benutzen.
- Wie kann man diese Veränderlichkeit verkleinern? (Normalisierung)
- Wie kann das System mit restlichen Unterschieden handeln? (Anpassung)

Die Herausforderung des Sprechers (II)

- Normalisierung Methoden: Zum Beispiel, Frauen sprechen in einer hoher Frequenz als Männer („pitch“). Die Normalisierung befreit die Stimme von diese Eigenschaft.
- Akustische Anpassung: Verbmobil hat man mit 25 Sprachestunden trainiert, aber es muss einen Algorithmus haben, der schätzt die Parameter in nur 30 Sekunden ein.
- Lexikalische Anpassung: Einige Variationen des Sprechers kann man nicht im Akustischen Level lösen. Zum Beispiel, man spricht das gleiches Wort in verschiedenen Arten aus („Ausssprache“).
- Sprechen Steuer: Man zieht in Betracht das Sprechen Steuer um die richtige, deutliche Anpassungstechnik zu anwenden.

Schnelle Suche für große Wortschatz Spracherkennung

- Verbmobil muss „Sprache-Sprache“ in fast Realzeit übersetzen.
- Die folgenden Module sind abhängig von dem Ergebnis der Erkennung.
- Aber es gibt Module, die arbeiten mit teilweise Ergebnisse.
- Diese Module haben eine wachsende Erkennungsausgabe.
- Es gibt die folgenden Methoden:
 - Beschleunigung Methoden um Innen-Wort Erkennung.
 - Suche und Beschleunigung Methoden um durch-Wort Erkennung.
 - Beschleunigung Methoden um stimmhafte Erkennung.
 - Wachsende Aufbereitung Methoden um die Reaktionszeit zu verkleinern.

Der „prosodic“ Modul (I)

- Verbmobil benutzt „prosodische“ Information.
- „Prosodische“ Information: Sprachsegments länger als ein Phonem: Silben, Wörter, Sätze und ganzen Turnus.
- Man benutzt Eigenschaften wie: „picht“, Sprechen Steuer, Qualitätssprache, Dauer, Pausen, Rhythmus, usw.
- „Prosodische“ Funktionen benutzt Verbmobil:
 - Abgrenzungen
 - Akzente
 - Satzstimmung
 - emotionaler Zustand

Der „prosodic“ Modul (II)

- Abgrenzungen
 - Fünfter geht bei mir, nicht aber neunzehnter.
 - Fünfter geht bei mir nicht, aber neunzehnter.
 - *The fifth is possible for me, but not the nineteenth.*
 - *The fifth is not possible for me, but the nineteenth would be OK.*
- Betonung
 - Ich fahre doch am Montag nach Hamburg.
 - Ich fahre DOCH am Montag nach Hamburg.
 - *I will go on Monday to Hamburg.*
 - *I will go on Monday to Hamburg after all.*

Der „prosodic“ Modul (III)

- Satzstimmung
 - Treffen wir uns bei Ihnen?
 - Treffen wir uns bei Ihnen!
 - *Do we meet at your place?*
 - *Let us meet t your place!*
- Abgrenzungen und Satzstimmung
 - Dann machen wir das vielleicht. Ab dem sechsten geht das.
 - Dann machen wir das vielleicht. Ab dem sechsten? Geht das?
 - *Perhaps, we should do that. It is possoble after the sixth.*
 - *Then let's do it that way. Maybe after sixth? Is that possible?*

Warum ist Prosodie benutzen schwierig?

- Wie viele „prosodic“ Klassen gibt es? (zwei, drei oder mehr Abgrenzungen in einem Satz)
- Segmentär (Wörterverkettung) und Supersegmentär (Prosodie) beeinflussen gegenseitig.
- Die verschiedenen „Prosodische“ Funktionen mischen sich gegenseitig ein.
- Es gibt eine Verbindung zwischen die „Prosodische“ Parameter (ein kleiner Wert von einem Parameter kann einen grossen Wert von anderem Parameter entschädigen).
- Die Benutzung des „Prosodisches“ Mittels ist wahlweise.
- Die Benutzung der Prosodie ist Sprecher- und Sprache-spezifisch.

Transfer und Übersetzung

Transferkonzepte für Vermobil (I)

Transfer

wird als allgemeiner Begriff für die Modellierung von Übersetzungsrelationen gebraucht.

Maschinelle Übersetzungssysteme waren früher erfolgreich nur für geschriebene Sprache entwickelt worden. Die innerhalb dieser Systeme verfolgten Transferkonzepte orientieren sich an den Charakteristika der geschriebene Sprache.

Transferkonzepte für Vermobil (II)

Übersetzung gesprochener Spontansprache

Spontansprache ist frei formulierte Alltagssprache, bei der ein Sprecher nicht etwa vorbereitete Texte vorliest. Gedankengänge werden fortlaufend in Sprache umgesetzt, wobei sehr häufig auch ungrammatische Sätze entstehen. Vermobil muß deshalb mit abgebrochenen Sätzen, Selbstkorrekturen umgehen können.

Transferkonzepte für Vermobil (III)

Die folgende Punkte haben große Bedeutung für die Entwicklung eines Transferkonzepts für Vermobil:

- die Dolmetschfunktionen, für Vermobil ausgelegt werden soll,
- die Tatsache, dass Vermobil spontan gesprochene Sprache als Input für die Übersetzung verarbeiten soll ,
- Die konstitutiven Merkmale des Diskurstyp für für den Vermobil ausgelegt werden soll.

Transferkonzepte für Vermobil (IV)

- Welche neuen Konzepte müssen für die Modellierung von Übersetzungsrelationen dialogisch gesprochene Sprache entwickelt werden?
- Welche Schnittstellen der Transferkomponente zu anderen Modulen sind vorzusehen?
- Welche Anforderungen stellt die Transferkomponente als Input?
- Welchen Output stellt die Transferkomponente für die Generierung zur Verfügung?
- Welche Anforderungen stellt die Entwicklung effizienter Transferstrategien an den Basisformalismus?

Die Dolmetschfunktion für Vermobil (I)

Für Vermobil sind zwei verschiedene Grundfunktionen für das Dolmetschgerät definiert, die unterschiedliche Konsequenzen haben.

- Volle Gesprächsübersetzung
von der Muttersprache der Gesprächsteilnehmer (Deutsch, Japanisch) in die gemeinsame Verhandlungssprache Englisch. Es wird aber vorausgesetzt, dass Englisch von den Gesprächspartner passiv mindestens beherrscht wird. (Passiv, d.h. dass sie dem Gespräch problemlos folgen können)
Bei dieser Funktion werden abgeschlossene Gesprächsabschnitte übersetzt. Für die Übersetzung steht die vollständige Analyse von strukturell und inhaltlich abgeschlossenen Äußerungseinheiten zur Verfügung

Die Dolmetschfunktion für Vermobil (II)

- Gesprächsunterstützende Übersetzung
aus der Muttersprache in die gemeinsame Verhandlungssprache, wenn das Gespräch fortlaufend in der Verhandlungssprache geführt wird und die Gesprächsteilnehmer an einem bestimmten Punkt des Gesprächs Formulierungsschwierigkeiten haben. Hier bezieht sich Transfer auf Einheiten, die eingebettet sind in Einheiten der Zielsprache.

Es ist hier wichtig, dass ein Gesprächspartner mitten in einer Äußerung von der Verhandlungssprache in die Muttersprache nicht wechseln kann.

Diese Funktion setzt also eine Analyse des auf Englisch geführten Gesprächs voraus.

Bestimmung von Transfereinheiten eines Dialoges (I)

Transfer in geschriebenen Sprache geht von Sätzen aus. Man kann also voraussetzen, dass die Sätze grammatisch korrekt sind.

Diese Voraussetzung kann für die Übersetzung gesprochener Sprache nicht gemacht werden. Folgende Phänomene sind bei der Modellierung von Übersetzungsrelationen

In Vermobil zu berücksichtigen:

- Äußerungen der gesprochenen Sprache sind doppelt strukturiert. Primäre Gliederungsebene ist nicht Einteilung in Sätze (syntaktische Regeln), sondern in kommutative Sinneinheiten (Sprecherwechsel, Gliederungssignale der Rede z.B. Gliederungspartikel, Kontaktpartikel, Pausen)

Bestimmung von Transfereinheiten eines Dialoges

(II)

- Redepartikel: Der Sprecher modifiziert das Potenzial seiner Äußerung. Die Redepartikel haben eine relativ freie Stellung. Für ihre Interpretation und Identifikation müssen besondere Strategien der Analyse gefunden werden.
Diese Potential in der Zielsprache wird oft nicht durch gleiche Partikel sondern durch andere Sprachmittel realisiert (z.B. Wortstellung, Intonation).
- Korrekturen: Zu Gespräch gehören auch Korrekturen. Der Sprecher könnte Äußerungen oder Teile während Äußerungsakt korrigieren.

Beispiel (I)

Gesprochene Sprache kennt keine Interpunktion.
Betonung und Phrasierung ersetzen Punkt und Komma.

Die Wortfolge "**Ja-zur-Not-geht-es-auch-am-Samstag**"
kann je nach Betonung

als *Bestätigung des Termins* "Samstag" interpretiert werden:

„Ja, zur Not geht es auch am Samstag“

oder als eingeschränkte Annahme eines Termins mit Gegenvorschlag:

"Ja, zur Not! Geht es auch am Samstag?"

Beispiel (II)

Wenn der Sprecher sagt:

**"Ja, ich weil also würde mal sagen äh vorschlagen, wir könnten uns am
äh 7. treffen so im Mai,,**

, so würde dieser Satz von einem an der Schriftsprache orientierten System abgelehnt und der Sprecher müßte den Satz wiederholen.

Durch Kombination von statistischen und linguistischen Verfahren wird Verbmobil jedoch so fehlertolerant, dass der Dialogakt **"suggest_date"** mit der **Datumsangabe "7. Mai"** aus der Äußerung extrahiert und die Übersetzung

"How about the seventh of May?"

ausgegeben wird.

Beispiel (III)

Nur durch die Berücksichtigung der Akzent können Mehrdeutigkeiten auch von einzelnen Wörtern wie "noch" für die Übersetzung aufgelöst werden.
Lautet die Eingabe:

"Wir brauchen noch einen Termin"

ohne Akzent auf "noch", so übersetzt Verbmobil mit

"We still need a date".

Wird "noch" jedoch betont, so wählt Verbmobil aufgrund der anderen Satzbedeutung die Übersetzung

"We need another appointment".

Beispiel (IV)

Die Transferregeln von Verbmobil müssen in Sortentests auf Wissen zurückgreifen.

Das Wort **"vor"** in dem Satz **"Wir treffen uns vor dem Hotel"** durch **"in front of"** zu übersetzen,

aber bei der Eingabe **"Wir treffen uns vor der Tagung"** die Übersetzung **"before"** zu wählen.

Beispiel (V)

Die Übersetzung muß auch kontextabhängig erfolgen und den Dialogverlauf berücksichtigen. So muß der Satz

"Geht es bei Ihnen?"

von Verbmobil als

"Do we meet at your place?"

übersetzt werden, wenn vorher gefragt wurde

"Wo treffen wir uns?".

Dagegen lautet die korrekte Übersetzung der identischen Eingabe

"Is it possible for you?",

wenn vorher

"Sollen wir uns im April treffen?"

geäußert wurde.

Ein Handlungsschema

In dem VERBMOBIL maßgebenden Szenario sind alle Äußerungen innerhalb eines bestimmten Handlungsschema verankert.

Ein Handlungsschema ist charakterisiert, dass für Sprecher und Hörer ein Ziel vorgegeben ist, zu dem das Gespräch führen soll und dass es einen Ablauf vorstrukturiert, nach dem dieses Ziel verfolgt werden kann, so dass an jedem Punkt des Gesprächs festgelegt ist, welche Sprachhandlungen wechselseitig aufeinander folgen können.

Das Handlungsschema - Terminvereinbarung (I)

Das Ziel ist in Abstimmung der wechselseitigen Interessen einen Termin für ein Treff festzulegen.

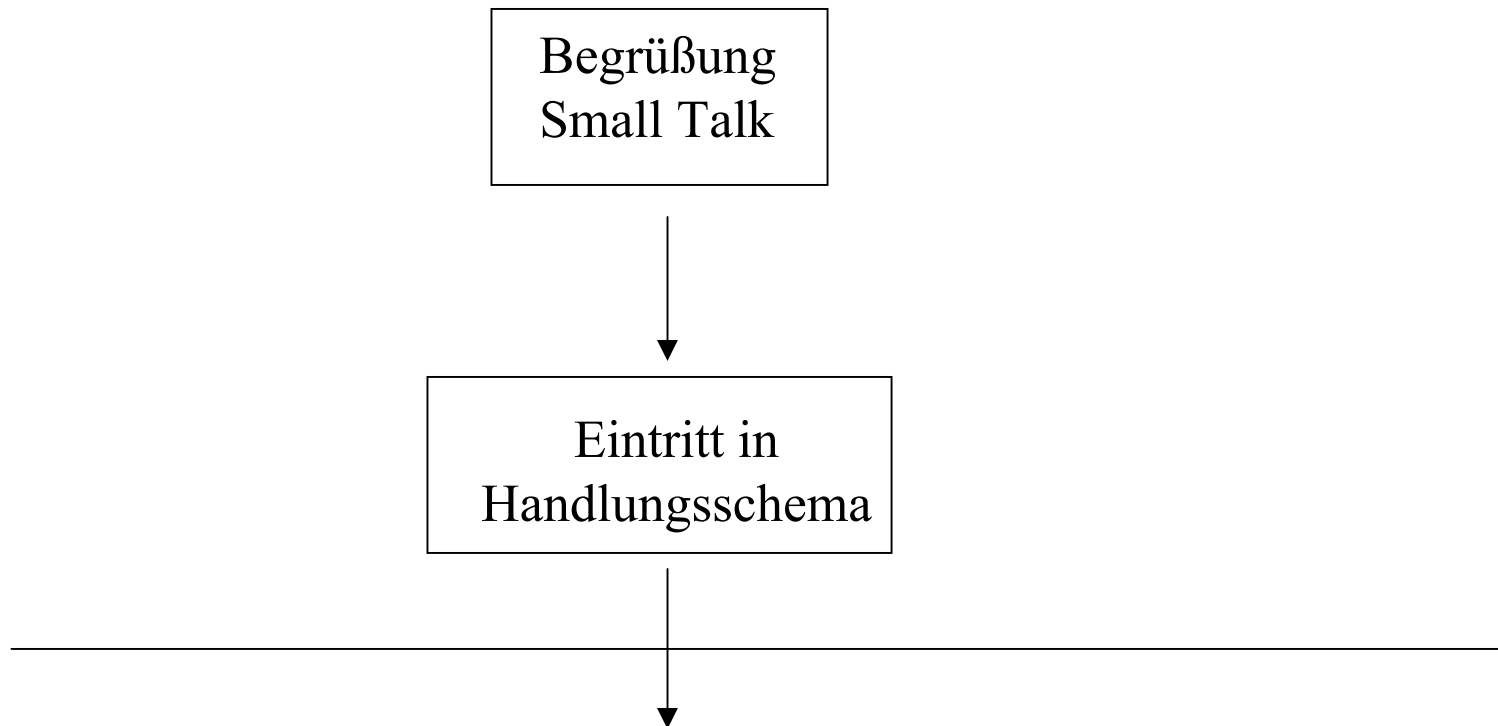
Ein Dialog ist beendet, wenn für alle Teilnehmer erkennbar das Dialogziel erreicht ist – Der Zeitpunkt und der Ort eines Treffens.

Wir haben hier 3 Typen von Handlungssequenz. Für jede Sequenz sind bestimmte Sprachakte typisch.

- Einleitungssequenz: hier wird der Kontakt zwischen den Dialogteilnehmer und das Gesprächsziel festgelegt.
- Vereinbarungssequenz: Vereinbarung des Termin (obligatorisch) , des Ortes (fakultativ).
- 3. Schlußsequenz: das Ergebnis wird hier nochmal festgehalten.

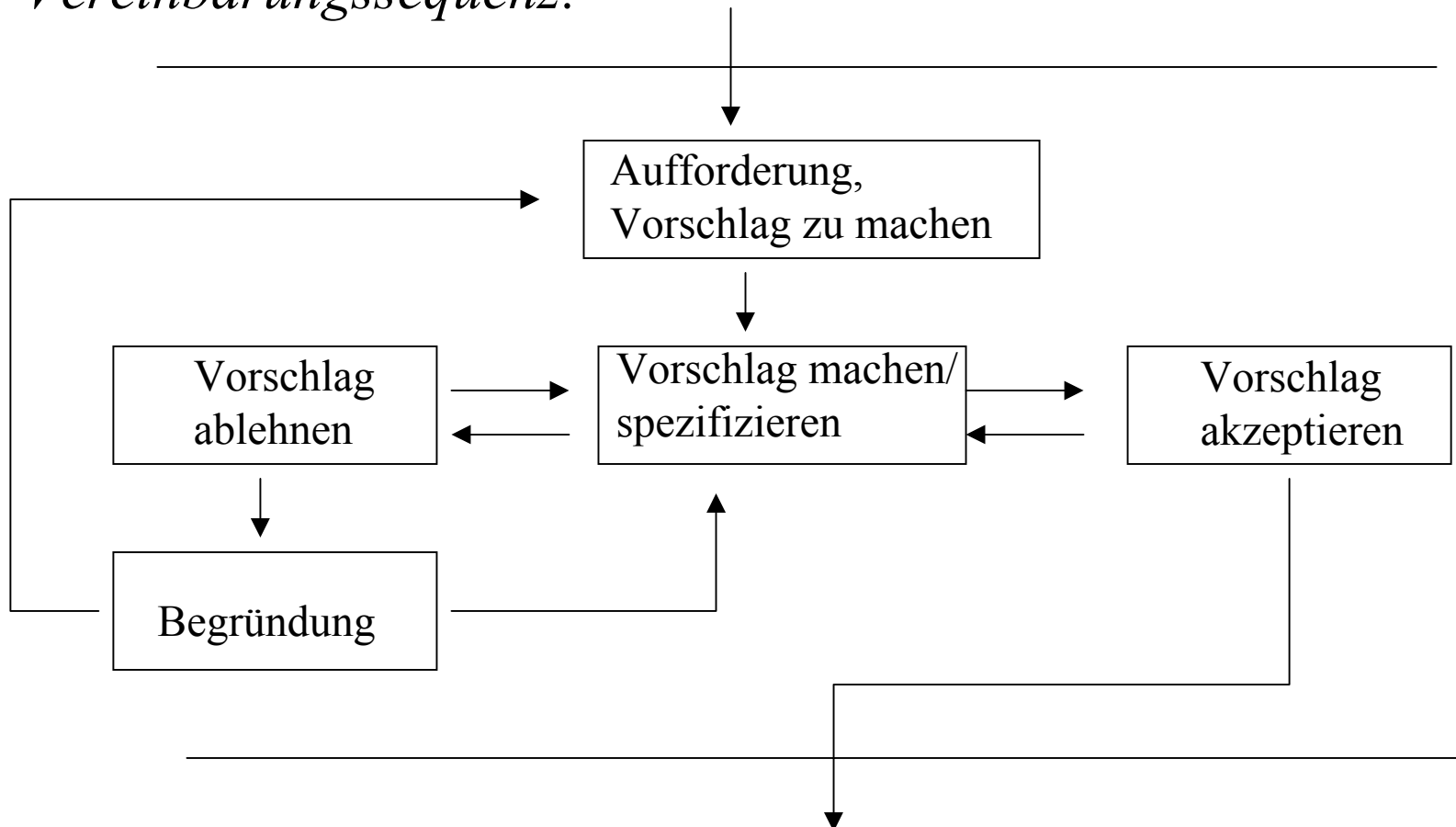
Das Handlungsschema - Terminvereinbarung (II)

Einleitungssequenz:



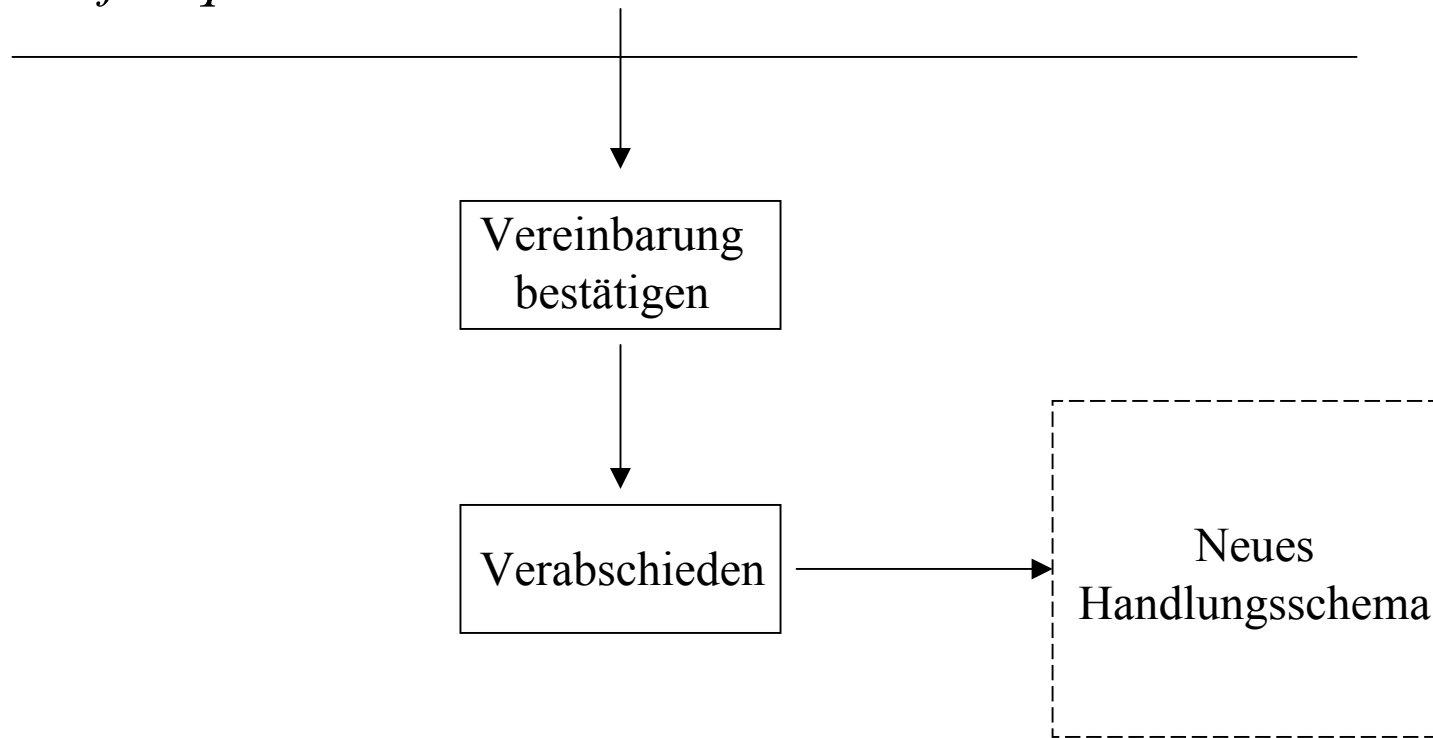
Das Handlungsschema - Terminvereinbarung (III)

Vereinbarungssequenz:



Das Handlungsschema - Terminvereinbarung (IV)

Schlußsequenz:



Transfer - eine Kontrollinstanz in Vermobil (I)

Alles bedeutet, dass der klassische Architekturansatz der maschinellen Übersetzung geschriebener Sprache nicht uneingeschränkt übernommen werden kann. Die einzelnen Komponenten können nicht sequentiell durchlaufen werden. Sie müssen kommunizieren und Information austauschen. Sie sollten Objekten (Sätze , Phrasen und Wörter) bearbeiten und sprachlichen Phänomenen (Abbrüche, Korrekturen, Wiederholung etc.) behandeln.

Die Einheit Satz ist ziemlich unproblematisch (klassische, maschinelle Übersetzung).

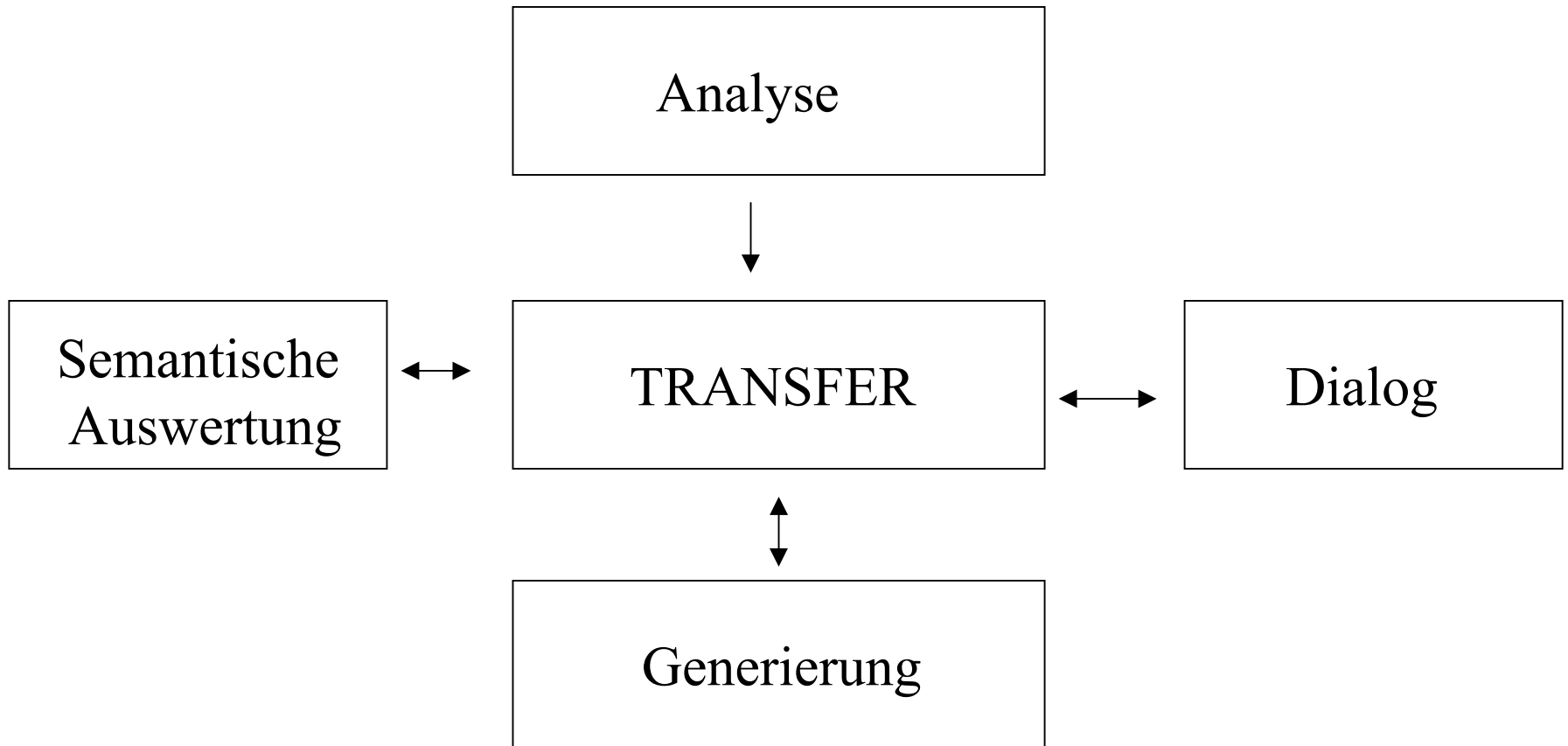
Die Einheiten Phrase und Wort müssen in einem Kontext bearbeitet werden. Abbrüche, Korrekturen verlangen die Tilgung.

Transfer - eine Kontrollinstanz in Vermobil (II)

Transfer wird nicht nur verstanden als die Komponente, die die eigentliche Übersetzung realisiert, dass sie für die Generierung notwendige Informationen erzeugt werden, sondern zusätzlich als eine Kontrollinstanz für Übersetzbarkeit.

Transfer erhält innerhalb der Gesamtsystemarchitektur eine Steuerfunktion für die Modulen zur semantische Auswertung, Dialogauswertung und zur Generierung.

Transfer – eine Kontrollinstanz in Vermobil (III)



Literaturverzeichnis

- Verbmobil: Foundations of Speech-to-Speech Translation, Wolfgang Wahlster, Springer Verlag, 2000, Seite 3-20 und 635-668
- Transfer in Verbmobil, F. Caroli, R. Nübel, B. Ripplinger, J. Schütz, Vm Report 11, Mai 1994
- The Verbmobil Domain Model Version 1.0, J.J. Quantz, M. Gehrke, U. Küssner, B. Schmitz, Vm Report 29, September 1994
- A Robust and Efficient Three-Layered Dialogue Component for a Speech-to-Speech Translation System, J. Alexandersson, E. Maier, N. Reithinger, Vm Report 50, Dezember 1994
- <http://verbmobil.dfki.de/verbmobil>