
Ein Vokabeltrainer auf der Grundlage von GermaNet und Mapa - Mapping Architecture for People's Associations

Abstract

Dieser Beitrag zeigt unser studentisches interuniversitäres Projekt¹ eines prototypischen Programms zur Konstruktion von „Wissensnetzen“. Dabei soll der Benutzer Wörter und andere Daten aller Art verlinken können, um sein Wissen auf eine kognitiv adäquate Art und Weise zu repräsentieren.

Als Beispiellapplikation zu diesem Projekt haben wir einen Vokabeltrainer konzipiert, der die Daten von GermaNet² (KUNZE 2001) als Netzstruktur darstellt, mit dessen Hilfe Lerner deutsche Wörter verstehen und lernen können.

1 Einführung

1.1 Projektrahmen

Das vom Projekt MiLCA unterstützte Studienprojekt Mapa³ (Mapping Architecture for People's Associations), fand an drei Universitäten statt.

Sieben Studenten des Studiengangs „Cognitive Science“ an der Universität Osnabrück, eine Studentin der Linguistik und der Sprachlehrforschung an der Universität Bochum und zwei Studentinnen des Studienganges „Allgemeine Sprachwissenschaft und Nebenfächer“ an der Universität Tübingen arbeiteten für dieses Projekt ein Jahr lang zusammen.

Im ersten Semester wurde in wöchentlichen Chats, im Whiteboard, in Telefonkonferenzen und per Email diskutiert und das Konzept des Gesamtprojekts und die Aufgabenverteilung spezifiziert. Die Aufgabe der Osnabrücker Studenten war dabei die Ausarbeitung des Frames von Mapa, die Tübinger und Bochumer Studentinnen erarbeiteten den Vokabeltrainer dazu. Im zweiten Semester folgte der Programmier-Teil

des Projekts. Die Koordination und Kommunikation erfolgte weiterhin über Telefon-Konferenzen und per Email, außerdem gab es einen gemeinsamen Workshop und zwei Programmierwochen, in denen sich alle Projektteilnehmer ausführlicher absprechen konnten.

Das Master-Projekt der Osnabrücker Studenten begann im Oktober 2002 und wird am 15. Oktober 2003 enden.

1.2 Konzept von Mapa

Das Mapa-Konzept wurde von den Studierenden selbst entwickelt⁴. Ziel dieses Projekts war ein Programm, mit dem Nutzer ihre eigenen Wissensnetze erstellen und darin Daten aller Art einbetten können, z.B. Bilder, Videos, Dokumente oder Emails.

Darüber hinaus soll es durch Kollaboration möglich sein, über das Internet Mapa-Netze von anderen Benutzern zu empfangen und das eigene Netz um deren Informationen zu erweitern. Gibt es in dem empfangenen Netz einen Knoten, über der Benutzer schon verfügt, wird dieser nicht neu erstellt. Seine Relationen werden dem schon existierenden Knoten hinzugefügt und so ins Netz aufgenommen.

1.2.1 Funktionalität

Mapa stellt ein Werkzeug zum „kognieren“ (überlegen, phantasieren, assoziieren, denken ...) dar, das möglichst viele verschiedene Ansprüche von Benutzern erfüllen soll. Deshalb können auch unterschiedliche Informationstypen (Text, Bilder, Videos, Emails...) integriert werden. Das Tool soll Benutzern ermöglichen, ihr Wissen so zu strukturieren, dass sie ihre Informationen am intuitivsten und für sie sinnvollsten

anordnen können. Ziel ist es, Wissen so anzuordnen, wie es möglicherweise auch im Gehirn angeordnet ist; das Programm versucht, die Arbeitsweise des Gehirns kognitiv adäquat wiederzugeben (VESTER 2002).

Der Benutzer soll möglichst viel Freiheit bei der Gestaltung haben. Statt Restriktionen, die die Möglichkeiten einengen, sollen Konventionen für Verständlichkeit sorgen. Wir gehen davon aus, dass der Benutzer nur sinnvolle Einträge macht, die er im Nachhinein rekonstruieren kann.

Das Netzwerk, das man mit Mapa konstruieren kann, hat nicht den Anspruch, Wissen zu repräsentieren sondern Stichworte für Wissen. Man muss also nicht das gesamte Wissen ins Netz schreiben. Es genügt, ein Wort oder einen Satz zu schreiben, durch den das Wissen zum Thema aktiviert wird.

Durch weltweite Vernetzung der Anwender und deren Kollaboration kann man das Wissen anderer Benutzer verwenden und sein eigenes Netz weiterentwickeln.

Jedes Jahr wird 1 Milliarde Gigabyte neuer Daten generiert. Durch die Zunahme verfügbaren Wissens wird es immer wichtiger, dies zusammenzufassen und zu strukturieren. Darüber hinaus müssen Möglichkeiten gefunden werden, das Wissen möglichst einfach zu erwerben und dann zu behalten (TERGAN 2003).

Als wesentliche Bausteine kognitiver Kompetenz gelten:

- Wissensmanagement
- möglichst schnelle Übersicht über viele Daten
- möglichst schnelle Analyse der relevanten Daten
- Verständnis komplexer Sachverhalte
- möglichst langes Behalten der Informationen

1.2.2 Visualisierung

Einer der Vorteile von Mapa ist, dass durch die graphische Zusammenfassung jedes Konzept nur einmal eingeführt wird. Anders als bei Texten,

wo jedes Konzept in jeder Aussage genannt werden muss, reicht es durch die Relationspfeile bei einer Graphik, jedes Konzept nur einmal aufzuschreiben.

Im Gegensatz zu Texten werden Zusammenhänge graphisch deutlicher. Bei mehreren Aussagen über mehrere Konzepte können diese verstreut im Gesamttext sein. Durch die Relationspfeile in einer Graphik dagegen sieht man stets auf einen Blick alle Relationen eines Konzepts.

Es ist auch wissenschaftlich nachgewiesen, dass graphische Darstellungen reichhaltigere Gedächtnisspuren hinterlassen als textuelle (PAIVIO 1971). Man kann Informationen also besser behalten, wenn man sie in einem Netz darstellt, als wenn man sie nur als Text aufschreibt.

Modell und Ansicht des Modells sind voneinander unabhängig und getrennt. Das bedeutet, man kann je nach momentaner Präferenz sein abstraktes Netz auf verschiedene Arten visuell dargestellt betrachten, als hyperbolischen Graphen oder als Touchgraphen⁵, siehe Abb. 1,2: Wichtig für optimales Lernen und eine gute Wissensorganisation ist es, Inhalte auf eine Weise darzustellen, die der mentalen Repräsentation dieser Inhalte entgegenkommt. Das muss notwendigerweise graphisch sein, da das Gehirn Bilder und Zusammenhänge abspeichert, keine „Texte“.

Die Vorteile einer graphischen Darstellung (Map) sind:

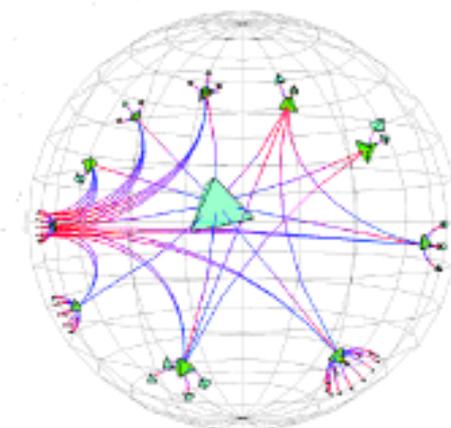


Abb. 1: hyperbolischer Graph

Vokabeltrainer

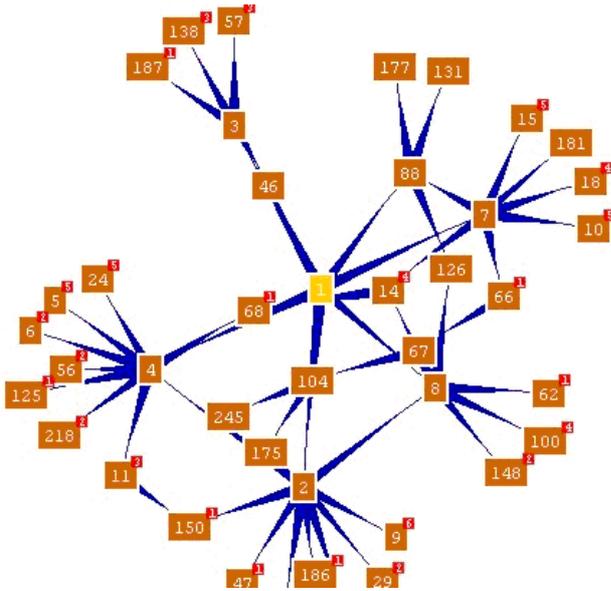


Abbildung 2: Touchgraph

- Man kann automatisch den Kontext einer Informationseinheit sehen, da er in ihrem Umfeld dargestellt ist.
- Man hat Freiheiten in der Reihenfolge der Exploration der Informationen einer Map.
- Mapping erfordert Verringerung der Komplexität, der Benutzer sieht automatisch die wichtigsten Punkte des Themas komprimiert.
- Der Benutzer bekommt sofort einen schnellen Überblick über den ihm unbekanntem Stoff.

Das von uns entwickelte Lern- und Organisationsstool besitzt folgende visuelle Features:

- Visualisierung einzelner Knoten in der Visualisierung des Netzes (Bilder, Videos etc, die ein Konzept verdeutlichen)
- Typisierung durch Farben/ Formen
- Zooming (man kann sich einen ausgesuchten Teil des Netzes vergrößert darstellen lassen)
- Chunking (man kann sich nur ausgewählte Knoten anzeigen lassen).

Große Vorteile dieses Konzepts gegenüber Mindmaps sind seine Dreidimensionalität, seine unhierarchische Oberflächenstruktur und die beliebig vielen Relationen, die ein Knoten haben kann. So können auch große Mengen an Knoten

(und ihren konzeptuellen Strukturen) anschaulich dargestellt werden.

1.2.3 Benutzbarkeit

Wir haben versucht, allgemeine Kriterien für die Benutzbarkeit von Software zu berücksichtigen (ISO 9241⁶):

- einfacher Aufbau und Transparenz, um die Anwendung des Programms zu erleichtern
- Skalierbarkeit in möglichst vielen Dimensionen
- leichte Erweiterbarkeit des Programms für persönliche Veränderungen
- Konventionen-Bildung und minimale Restriktivität für möglichst vielseitige Benutzbarkeit
- Unterstützung verteilter Datenhaltung (Kollaboration) für Kommunikation zwischen Benutzern, ohne einen Administrator und einen Server zu benötigen (Peer to Peer)

2 Der Vokabeltrainer

2.1 Konzept

Unser Vokabeltrainer basiert auf dem Konzept von Mapa und ist eine Beispielanwendung mit lexikalischen Daten. Er verwendet also die Ansatzmöglichkeiten und die Netzwerkstruktur von Mapa.

Die zugrunde liegende Idee ist, dass linguistisches Wissen, das ja semantisches Wissen beinhaltet, als Netzwerk von verknüpften Konzepten repräsentiert ist.

In das Netzwerk von Mapa werden Daten von GermaNet, das diese Informationen implementiert hat, eingebaut. Knoten sind Lemmata, Verbindungen sind semantische und morphologische Relationen wie Synonymie, Hyperonymie, Derivationen etc. Ein lexikalischer Eintrag wird als Knoten mit dem ihn umgebenden Feld dargestellt, siehe Abbildung 3.

Wir verwenden die XML-Version von GermaNet, in dem deutsche Wörter als Synsets⁷ (FELLBAUM 1998) kodiert und die Relationen zwischen den Wörtern und Konzepten gespeichert

chert sind. Mapa extrahiert diese Daten und speist sie in sein Netz-Schema ein. Sie sind die Basis des Vokabeltrainers.

Aus GermaNet importierte Relationen sind:

- Semantische Relationen: Hypo-/ Hyperonymie, Synonymie, Antonymie, Meronymie, Holonymie, Cause und Assoziation
- Lexikalische Relationen, z.B. Derivationen

Die Idee des Vokabeltrainers ist, dass in natürlicher Sprache Wörter, also Konzepte in semantischem Kontext auftreten. Aus der kognitiven Psychologie ist bekannt, dass im Gehirn Bedeutungen netzwerkartig abgespeichert werden. Und aus der kognitiven Informationsverarbeitung wissen wir, dass semantische Verarbeitung die Erinnerungsleistung verbessern kann (VESTER 2002). Folglich müsste sich auch der Fremdspracherwerb verbessern, wenn Vokabeln im Kontext gelernt werden. Wir versuchen also, dem Vokabeltrainer ein möglichst kognitiv adäquates Modell der Sprache zugrunde zu legen.

Die Zielgruppe dieses Vokabeltrainers sind Lerner mit einem gewissen Vorwissen an Wortschatz und Grammatik, die ihr Vokabelwissen vertiefen möchten. Ein Vorteil dieses Vokabeltrainers ist, dass er einsprachig deutsch ist. Weder zum Verständnis der Wörter noch zur Abfrage sind Übersetzungen in eine andere Sprache notwendig, sodass die Anwender den Vokabeltrainer ohne Vorkenntnisse anderer Fremdsprachen benutzen können.

Da bei dieser Art von Vokabeltrainer die Bedeutung eines Wortes vorerst nur aus dem Kontext seiner Nachbarwörter und der Relationen zwischen ihnen erschlossen werden kann, muss der Lerner notwendigerweise schon Wörter kennen, mit deren Hilfe er die Bedeutung von Nachbarwörtern verstehen kann. Außerdem hat der Vokabeltrainer keine Komponenten, mit denen man Grammatik und Satzbau trainieren könn-

te, da wir uns auf den Netzwerkaspekt des Wortschatzes konzentriert haben.

2.2 Surf-Modus

Um sich im Netz zu bewegen, wählt der Lerner ein für ihn interessantes Einstiegswort aus, das ihm mit seinen Nachbarwörtern präsentiert wird. Von dort kann er eine Sequenz von Knoten, also Wörtern im Graphen, besichtigen und sich von Nachbarwort zu Nachbarwort klicken. Die Pfade, die ihm zur Verfügung stehen, sind die semantischen Relationen, selbstverständlich kann er aber auch zu einem ganz anderen Wort springen.

Diese Methode der Exploration bietet zwei Vorteile: Zum einen wird eine Kohärenz in der Abfolge der besuchten Wörter sichergestellt, da der Lerner sich ja die Wörter zu einem bestimmten Thema anschaut, das ihn gerade interessiert und kein Wort unabhängig von seinem Nachbarwort ist. Und zum anderen prägt sich der Lerner die Wörter schon ein, während er das Netz exploriert, weil er jedes Wort semantisch verarbeiten muss, um seine Bedeutung zu verstehen.

2.2.1 Auswahl von zu lernenden Vokabeln

Während der Lerner durch das Netz „surft“, kann er durch einfachen Mausklick Wörter, die

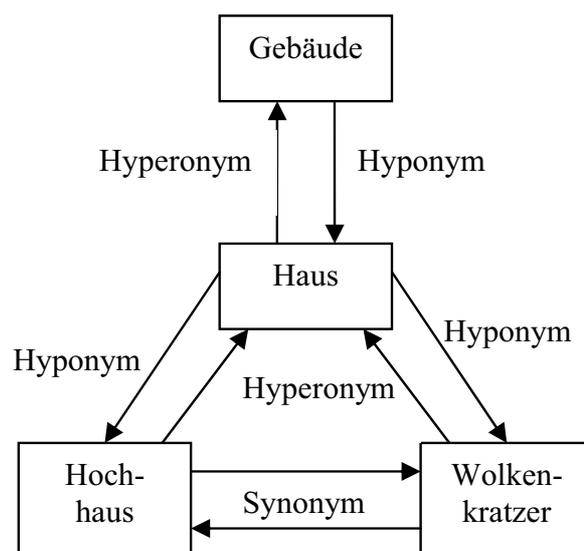


Abbildung 3: Knoten „Haus“ mit Nachbarknoten

er lernen möchte, in seine Vokabelliste aufnehmen. Dabei hat er die Wahl zwischen zwei Arten der Vokabelauswahl.

Er kann jedes Wort einzeln auswählen, um eine kohärente Lernliste zu bekommen und nur explizit gewünschte Wörter in seine Lernliste aufzunehmen.

Er kann aber auch automatisch alle Nachbarwörter eines bestimmten Wortes mit zum Lernen auswählen, die 1, 2, ... Relationen von ihm entfernt sind. Die vom Anwender auszuwählende „Suchtiefe“ n bestimmt, dass alle Wörter, die maximal n Relationen vom zentralen Wort entfernt sind, ausgewählt werden.

Im besten Fall erreicht er damit nur semantisch verwandte Wörter und kann schnell eine große Menge an Wörtern, ähnlich Wortfeldern, auswählen, z.B. zu einem bestimmten Stichwort wie „Bahnhof“, „essen“ etc. Es besteht allerdings schon bei relativ geringer Suchtiefe die Gefahr, auch semantisch unabhängige Wörter mit auszuwählen, z.B. Student \rightarrow Vorlesung \rightarrow Tafel \rightarrow Kreide \rightarrow weiß (Suchtiefe 4).

2.2.2 Individuelle Organisation

Lerner können eigene Lektionen anlegen, in die sie Vokabeln abspeichern und die sie unabhängig von anderen Lektionen lernen können. Außerdem wird der Pfad, den ein Lerner beim Surfen benutzt hat, abgespeichert und kann von ihm wieder angeschaut werden.

2.2.3 Erweiterung des Netzes

Wie vollständig das semantische Netz ist, hängt von den Kenntnissen und Interessen des Lernalters ab. Er kann neue Wörter hinzufügen, z.B. Fachwörter, umgangssprachliche Wörter etc.

Manchmal können auch persönliche Lernhilfen, Eselsbrücken nützlich sein.

Des Weiteren gibt es die Möglichkeit, Meta-Informationen zu einem Wort, z.B. Beispielsätze, mit ihm zu verknüpfen.

2.3 Vokabel-Abfrage

Dieser Vokabeltrainer-Prototyp bietet drei Abfragearten an, die vollständig auf GermaNet und seiner Netz-Idee basieren: Zentrales Wort einfügen, Relationen einfügen und Nachbarwörter einfügen.

2.3.1 Zentrales Wort einfügen

Bei dieser Übungsart bekommt der Lerner alle Nachbarwörter eines gesuchten Wortes und ihre Relationen gezeigt. Wird beispielsweise „Haus“ gesucht, sieht er „Hyperonym: Gebäude, Hyponym: „Hochhaus“, Hyponym: Wolkenkratzer“, siehe Abbildung 4.

Diese Übung sollte die einfachste der drei Übungen sein. Der Lerner erhält hier viel Kontextinformation, er muss nicht alle Nachbarwörter verstehen. Manchmal kann er das gesuchte Wort sogar aus einem Nachbarwort ableiten, z.B. „Haus“ aus „Hochhaus“.

Ein Problem bei der Konzeption war hier, eine geeignete Abfragemethode zu finden für Wörter, die mit mehreren anderen Konzepten durch dieselbe Relation verbunden sind. Bei Abbildung 4 beispielsweise enthält jedes der beiden unteren Kästchen eine mögliche Antwort auf die Frage nach dem Hyponym von „Haus“. Wir haben das Problem gelöst, indem das Programm jedes Wort akzeptiert, das durch diese Relation mit dem gegebenen Wort verbunden ist, aber so lange nach einer anderen Lösung gefragt wird, bis der Lerner die momentan gesuchte Vokabel eingibt.

2.3.2 Relationen einfügen

Bei dieser Übungsart bekommt der Benutzer ein zentrales Wort mit all seinen Nachbarwörtern und muss die Relationsarten zwischen ihnen bestimmen. Diese Übungsart ist leichter als die nachfolgende Übung, da man alle Nachbarwörter zur Verfügung hat und aus bekannten Relationsarten andere erschließen kann, siehe Abbildung 5.

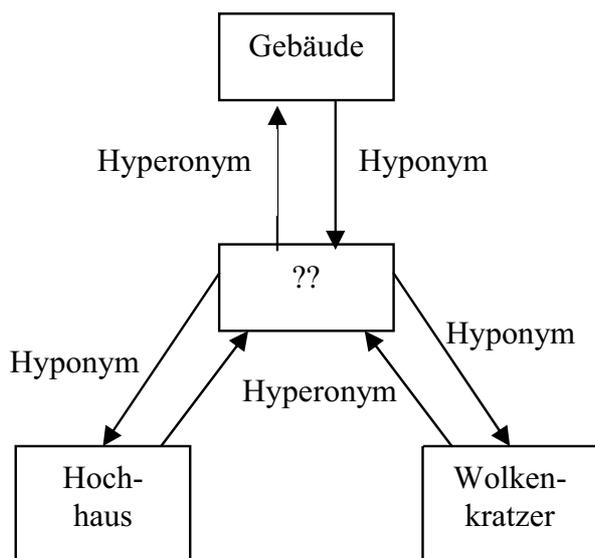


Abb. 4: Übungsmodus mit fehlendem zentralen Knoten

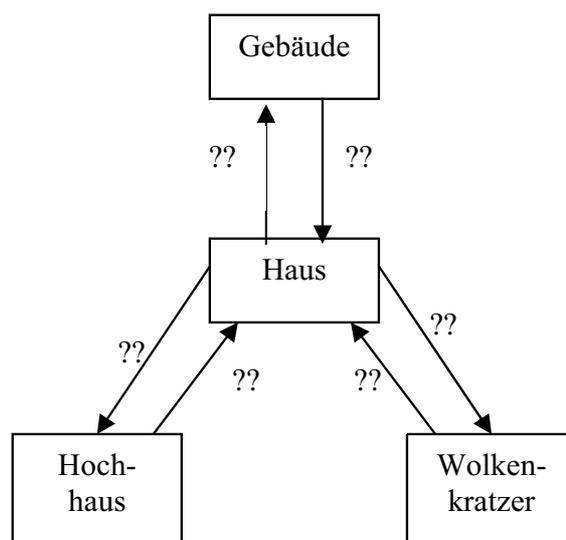


Abb. 5: Übungsmodus mit fehlenden Relationen

2.3.3 Nachbarwörter einfügen

Hier sieht der Lerner das zentrale Wort mit seinen Relationen, siehe Abbildung 6. Er muss nun zu jeder Relation das passende Nachbarwort eingeben. Gibt es mehrere Nachbarwörter, die über die gleiche Relationsart mit dem zentralen Wort verbunden sind (z.B. Synonym von „teuer“ = „wertvoll“, Synonym von „teuer“ = „kostspielig“), so spielt die Reihenfolge der Eingabe natürlich keine Rolle.

Genau wie bei der Übungsart mit fehlenden Relationen gibt es hier das Problem der Definition, wann der Lerner eine Vokabel gewusst hat und wann nicht. Mangels psychologischer Kenntnisse zu diesem Thema haben wir entschieden, vorerst zu definieren, dass der Lerner ein Wort gewusst hat, wenn er

- bei einem einzigen gesuchten Nachbarwort dieses und
- bei mehreren gesuchten Nachbarwörtern mindestens zwei von ihnen richtig benennen kann.

2.3.4 Ablauf der Vokabel-Abfrage

Für jedes gesuchte Wort bekommt der Lerner Informationen zu seinem Verständnis dargestellt.

Er nennt daraufhin seinen Lösungsvorschlag. Ist seine Antwort falsch, bekommt er noch einen zweiten Versuch, um sich selbst zu korrigieren und Alternativen auszuprobieren. Findet er beim zweiten Versuch die korrekte Lösung, wird das trotzdem als falsch abgespeichert, da er das Wort auf Anhieb wissen soll.

Hat er nur Groß- und Kleinschreibung nicht beachtet, wird er darauf hingewiesen. Wurde der Lerner alle momentan zu lernenden Vokabeln einmal abgefragt, so werden sie neu sortiert. Das heißt, Wörter, die jetzt als gelernt gelten, werden aussortiert, die verbleibenden werden, je nachdem, ob der Lerner sie gewusst hat oder nicht, in eine neue Reihenfolge nach Priorität der Abfrage sortiert, und der Lerner kann sich weiter abfragen lassen, so lange, bis er das Programm abbricht oder alle Wörter gelernt hat.

Am Ende der Sitzung bekommt er dann ein Feedback über seine Leistung:

- Anzahl der abgefragten Wörter
- Anteil der Wörter, die er bei dieser Sitzung richtig beantwortet hat

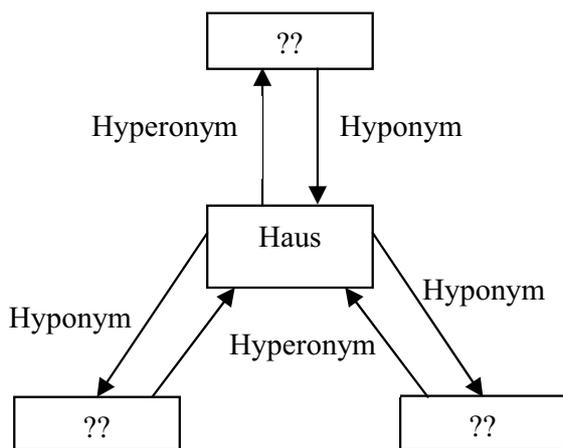


Abb.6 Übungsmodus mit fehlenden Nachbarknoten

- Anteil der Wörter, die er so oft in Folge richtig beantwortet hat, dass sie als gelernt gelten und nicht mehr abgefragt werden
- Gesamtanzahl der Wörter in der momentan abgefragten Lektion
- Anteil der schon vollständig gelernten Wörter dieser Lektion

So hat der Lerner einerseits ein Feedback über seine Leistung der letzten Abfrage und sieht, ob sein Lernen ausreicht, oder ob er mehr lernen sollte.

Außerdem bekommt er auf Wunsch ein Feedback in Bezug auf die gelernte Lektion. Er kann abschätzen, wie viel Arbeit wohl noch nötig sein wird, bis er die Lektion vollständig gelernt haben wird.

Wir glauben, dass die Selbsteinschätzung, die dem Lerner durch dieses Feedback ermöglicht wird, seine Motivation, Vokabeln zu lernen, stark fördert im Vergleich zu Vokabeltrainern ohne Feedback

2.3.5 Problem der nicht verbundenen Knoten

Bei der Vokabelabfrage nur auf der Grundlage von GermaNet können mit dem derzeitigen Programm Wörter, die der Lerner selbst eingefügt hat, und die keine Verbindung zu Nachbarknoten haben, nicht abgefragt werden. Man

könnte hier beispielsweise anbieten, mit Hilfe eines Übersetzungsprogramms die englische oder französische Übersetzung des Wortes zu suchen und mit dem deutschen Begriff zu verknüpfen.

2.4 Benutzer-Profil

Für jeden Benutzer des Vokabeltrainers wird ein persönliches Profil angelegt.

Es speichert Informationen über die Abfragesitzungen, das Lernverhalten des Benutzers und über alle besuchten Wörter, also über alle Wörter, über die der Lerner schon einmal „gesurft“ ist.

Da der Vokabeltrainer auf dem Karteikastenprinzip beruht, orientiert sich die Benutzermodellierung an diesem Prinzip.

Ein solcher Vokabel-Karteikasten besteht aus 6 Fächern (LEITNER 1977); dabei kommen noch nie gewusste Wörter ins 1. „Fach“, einmal gewusste ins 2. „Fach“ usw. bis zum letzten „Fach“. Diese Wörter gelten als gelernt und werden nicht mehr abgefragt. Wird ein Wort einmal nicht gewusst, kommt es unabhängig von dem „Fach“, in dem es vorher war, wieder ins 1. „Fach“ zurück.

Ein Wort, das man erst selten richtig beantwortet hat, vergisst man schneller als ein Wort, das man schon als richtige Lösung eingegeben hat, und es muss früher wieder abgefragt werden. So wird eine Abfragereihenfolge erstellt nach der Dringlichkeit der erneuten Abfrage. Diese Methode bietet den Vorteil, dass Vokabeln so spät wie möglich abgefragt werden können, kurz bevor der Lerner sie evtl. wieder vergessen würde. Wir erwarten, dass sich dadurch die Anzahl der momentan abzufragenden Vokabeln und die Abfragezeit verringern, vorausgesetzt, man verwendet den Vokabeltrainer in genügend kurzen Abständen. So kann sich der Lerner auf die in dem Moment relevanten Vokabeln konzentrieren.

Hat der Lerner schon alle Wörter einer Lektion gelernt, wird ihm das gemeldet. Er kann sie aber noch einmal lernen, wenn er beispielsweise sein Wissen wieder auffrischen will. Dann kom-

men alle Wörter zurück ins 1. „Fach“ des „Karteikastens“.

2.4.1 Vokabel-Profil

Im Benutzerprofil werden folgende Daten zu jeder Vokabel gespeichert:

- „Fach“ des „Karteikastens“, in dem es sich befindet
- Datum der letzten Abfrage
- Ob es bei der letzten Abfrage gewusst wurde oder nicht

Aus dem Datum der letzten Abfrage eines Wortes und seinem „Karteikasten-Fach“ wird berechnet, wann das Wort zum nächsten Mal abgefragt werden muss.

2.4.2 Surf-Profil

Das Benutzer-Profil speichert außerdem:

- alle besuchten Pfade (Wörter in der besuchten Reihenfolge mit Datum)
- Anzahl der besuchten Wortarten
- Anzahl der besuchten Relationsarten

Dadurch kann einerseits der Lerner alle abgespeicherten Pfade noch einmal anschauen, um seine Erinnerungen an den Wortschatz zu vertiefen oder z.B. ein Wort wieder zu finden. Diese Daten liefern aber auch Aufschluss über das Verhalten und die Kenntnisse des Benutzers. Man kann daraus eine Statistik der bevorzugten Wort- und Relationsarten ableiten, und evtl. auch Probleme erkennen. Wenn ein Lerner beispielsweise gewisse Relationsarten kaum benutzt, könnte das bedeuten, dass er ihre Bedeutung nicht verstanden hat.

2.4.3 Vokabellern-Statistik

Die dritte Art gespeicherter Daten ist für das Feedback des Lerners gedacht und enthält Informationen über sein Abfrage- und Lernverhalten:

- Anzahl der gelernten Wörter
- Anzahl der noch nicht gelernten, aber schon mehrmals gewussten Wörter
- Anzahl der gar nicht gewussten Wörter
- Anzahl der bei der letzten Abfrage gewussten Wörter
- Anzahl der Wörter, die in den nächsten Tagen wiederholt werden müssen

Diese Daten sind als Feedback für den Lerner gedacht, wie oben erklärt.

Die letzte Information ist besonders nützlich, weil sie den Lerner rechtzeitig daran erinnert, Vokabeln zu lernen, bevor er die schon halb gelernten Wörter wieder vergessen hat und er zu viele Wörter wiederholen muss.

3 Ausblick

Wir haben viele Erweiterungsvorschläge für dieses Projekt, jedoch weder die Zeit noch die Software, finanziellen Mittel oder andere Ressourcen, sie hatten im Rahmen unseres Studentenprojekts in die Realität umzusetzen.

3.1 Fehleranalyse und Feedback

Ein wichtiger Vorschlag wäre, statt nur zwischen richtiger und falscher Lösung zu unterscheiden, eine detaillierte Fehleranalyse mit Unterscheidung zwischen Rechtschreibfehlern und semantisch falschen Wörtern. Sie sollte einerseits Verfahren des approximativen String Matching (BAEZA-YATES 1998) berücksichtigen können – um die Eingabe mit den Vokabeln des Netzes zu vergleichen-, andererseits auch das Fehlerverhalten beim Fremdsprachenlernen modellieren können (CHANIER 1992).

Zusätzlich könnte man ein angemessenes informatives und motivierendes Feedback geben, z.B. „Du hast dich vertippt.“ oder „Du hast das Antonym der gesuchten Lösung eingegeben.“

3.2 Vorgefertigte Lektionen

Man könnte auch relevante Teile des Netzes als Lektion abspeichern und dem Lerner anbieten, z.B. Wörter zu einem bestimmten Thema, oder alle Nachbarwörter einer bestimmten Tiefe zu einem zentralen Wort.

3.3 Kollokationen

Genauso wichtig für den korrekten Sprachgebrauch sind Kollokationen (LEWIS 2000), denn ohne ihre Kenntnis macht der Deutsch-Lerner im besten Fall Fehler, im schlimmsten Fall wird er aber missverstanden. Möchte beispielsweise ein Engländer „eine Rede machen“, so will er sie wahrscheinlich halten. Der Deutsche versteht spontan vielleicht aber, dass er eine Rede vorbereiten, also „machen“ will.

Die Problematik bei Kollokationen ist aber einerseits, dass es keine genaue Grenze zwischen ihnen und offenen syntagmatischen Relationen gibt, und andererseits, dass es sehr viele Kollokationen gibt. Da sie in GermaNet nicht mit aufgenommen sind, müssten sie alle einzeln von Hand eingetragen werden, was einen immensen Aufwand darstellen würde.

3.4 Mündliche Abfrage

Man könnte zusätzlich zur schriftlichen eine mündliche Abfrage mit den gleichen Übungsarten anbieten. Diese Abfrageart wäre besonders geeignet, wenn der Lerner in kurzer Zeit seine Vokabeln auffrischen will. Er müsste seine Antwort nur laut aussprechen und dann selbst mit der richtigen Lösung vergleichen. Hier wäre es sinnvoll, ein Hörbeispiel oder zumindest die phonetische Lautschrift anzuzeigen, sodass der Lerner auch die korrekte Aussprache lernen kann. Optimal wäre natürlich eine automatische Spracherkennung, die die Antwort des Lerners analysiert. Solange diese Technologie aber nicht wirklich funktioniert, sollte die Selbstkorrektur des Lerners nicht in sein Benutzerprofil einge-

hen, da ihre Korrektheit nicht vom Programm überprüft werden kann.

3.5 Detaillierte Abfrage

Mit dem momentanen Programm kann man die in GermaNet kodierten, bedeutungstragenden Wortarten wie Nomina, Verben und Adjektive lernen. Das bedeutet, man lernt vielleicht in kurzer Zeit, sich mehr recht als schlecht verständlich zu machen, aber man lernt nicht, grammatikalisch korrekt zu formulieren. Für Lerner, die daran interessiert sind, muss ein zusätzlicher Modus angeboten werden, in dem sie grammatische Details der Wörter lernen. Das wären beispielsweise bei Nomina Artikel und Deklination, bei Adjektiven Flexion und bei Verben die Konjugation.

3.6 Einbindung von anderen Ressourcen

Außer der Netz-Struktur von GermaNet gäbe es noch viele andere Arten, die Bedeutung eines Wortes zu zeigen, z.B. über Bilder und Videos. So könnte man ein Nomen durch (ein) Bild(er) und ein Adjektiv durch Bilder der Gegenteile (z.B. eine große Frau vs. eine kleine Frau) darstellen. Verben könnte man mit einem Video zum Vorgang erklären.

Wort-Definitionen, wie man sie in einsprachigen Wörterbüchern findet, stellen eine andere, schwerer verständliche Art dar, ein Wort zu erklären. Der Lerner könnte dabei aber andere Wörter, mit denen es häufig auftritt oder Kontexte, in denen es verwendet wird, lernen und zur Sprachproduktion verwenden.

Zusätzlich wäre ein Übersetzungsprogramm hilfreich, um auch abstrakte Wörter zu verstehen.

3.7 Automatische Auswahl des Übungsprogramms

Mit all den oben vorgeschlagenen Erweiterungen könnte man die Vokabelabfrage so erweitern, dass das Programm je nach Lerngrad einer

Vokabel im Karteikastensystem einen anderen Übungsmodus auswählt.

Das hätte zum einen den Vorteil, dass durch verschiedene Hinweise auf das gesuchte Wort (lexikalisch durch verwandte Wörter oder semantisch durch Bilder, Videos) beim Lerner verschiedene Assoziationen geweckt werden und sich das Wort besser einprägen kann. Zum anderen gäbe es aber sicher einen großen pädagogischen Vorteil. Durch eine abwechslungsreich variierte Abfrage würden die Aufmerksamkeit, die Motivation und das Interesse am Vokabellernen vermutlich signifikant erhöht.

Anmerkungen

- ¹ Das diesem Projekt zugrunde liegende Projekt MiLCA (Medienintensive Lehrinhalte in der Computerlinguistik-Ausbildung) wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01 NM 167 A gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei der Autorin.
- ² Die in Tübingen erstellte deutsche Version von WordNet: <http://www.sfs.uni-tuebingen.de/lsd/> [Zugriff April 2004].
- ³ Mapa (spanisch: Landkarte): strukturierte Visualisierung von Wissensinhalten. Universität Osnabrück, Institut für Kognitionswissenschaft, Studentisches Projekt Mapa, <http://www.cogsci.uni-osnabrueck.de/-mapa>, mapa@cogsci.uni-osnabrueck.de [Zugriff April 2004].
- ⁴ Die Tübinger Dozentin Dr. Karin KRÜGER-THIELMANN und die Osnabrücker Dozenten Dr. Petra LUDEWIG, Dr. Claus ROLLINGER und Dr. Veit REUER betreuten das Projekt (REUER 2004).
- ⁵ TOUCHGRAPH LLC (2004). TouchGraph Website. <http://www.touchgraph.com> [accessed April 2004].
- ⁶ INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO) (2004). ISO Online Website, <http://www.iso.ch> [accessed April 2004].

- ⁷ Ein Synset ist ein aus synonymen Wörtern bestehender Knoten, der die Bedeutung dieser Wörter beschreibt.

Literatur

- BAEZA-YATES, R., NAVARRO, G. (1998). Fast Approximate String Matching in a Dictionary. University of Chile, Santiago de Chile, Dept. of Computer Science, <http://citeseer.nj.nec.com/1593.html> [Zugriff April 2004].
- CHANIER, T.; PENGELLY, M.; TWIDALE, M.; SELF, J. (1992). "Conceptual Modeling in Error Analysis in Computer-Assisted Language Learning Systems." In: SWARTZ, M. L.; YAZDANI, M. (eds.) (1992). Intelligent Tutoring Systems for Foreign Language Learning. Berlin et al.: Springer.
- FELLBAUM, CH. (ed.) (1998). WordNet – An Electronic Lexical Database. Language, Speech, and Communication. Cambridge, MA / London: MIT Press.
- KUNZE, C. (2001). „Lexikalisch-semantische Wortnetze." In: Carstensen, K.-U. et al. (eds.) (2001). Computerlinguistik und Sprachtechnologie: eine Einführung. Heidelberg; Berlin: Spektrum, Akademischer Verlag, 386-393.
- LEITNER, S. (1977). So lernt man richtig. Freiburg: Herder KG.
- LEWIS, M. (2000). Teaching Collocation: Further Developments in the Lexical Approach. Hove, England: LTP.
- PAIVIO, A. (1971). Imagery and Verbal Processes. New York: Holt, Rinehart & Fisher.
- REUER, V. ET AL. (2003). „Studienprojekte in den Bereichen Computerlinguistik und Cognitive Science." In: Sprache und Datenverarbeitung Heft 27(1/2) (2003), 185-202.
- TERGAN, S.-O. (2003). „Lernen und Wissensmanagement mit Hypermedien." In: Unterrichtswissenschaft 31 (4), 334-358.
- VESTER, F. (2002). Die Kunst, vernetzt zu denken. Ideen und Werkzeuge für einen neuen Umgang mit Komplexität. Der neue Bericht an den Club of Rome. München: DTV.