

# **Implementation einer Spracherkennung zur Steuerung von Audio-Software mit Hilfe des MIDI-Protokolls**

**von: Christian Neyens  
BAP1010**

- **Einführung**
- **Aufbau des Systems**
- **Funktionsweises**
- **Optimierung**
- **Umfrage**
- **Resultate**

# **Einführung**

**Sprache:**

**Erkennung und Verarbeitung**

**2 Arten von Systemen:**

- **Dictations Systems**

- **Command And Control**

# Einführung

## MIDI

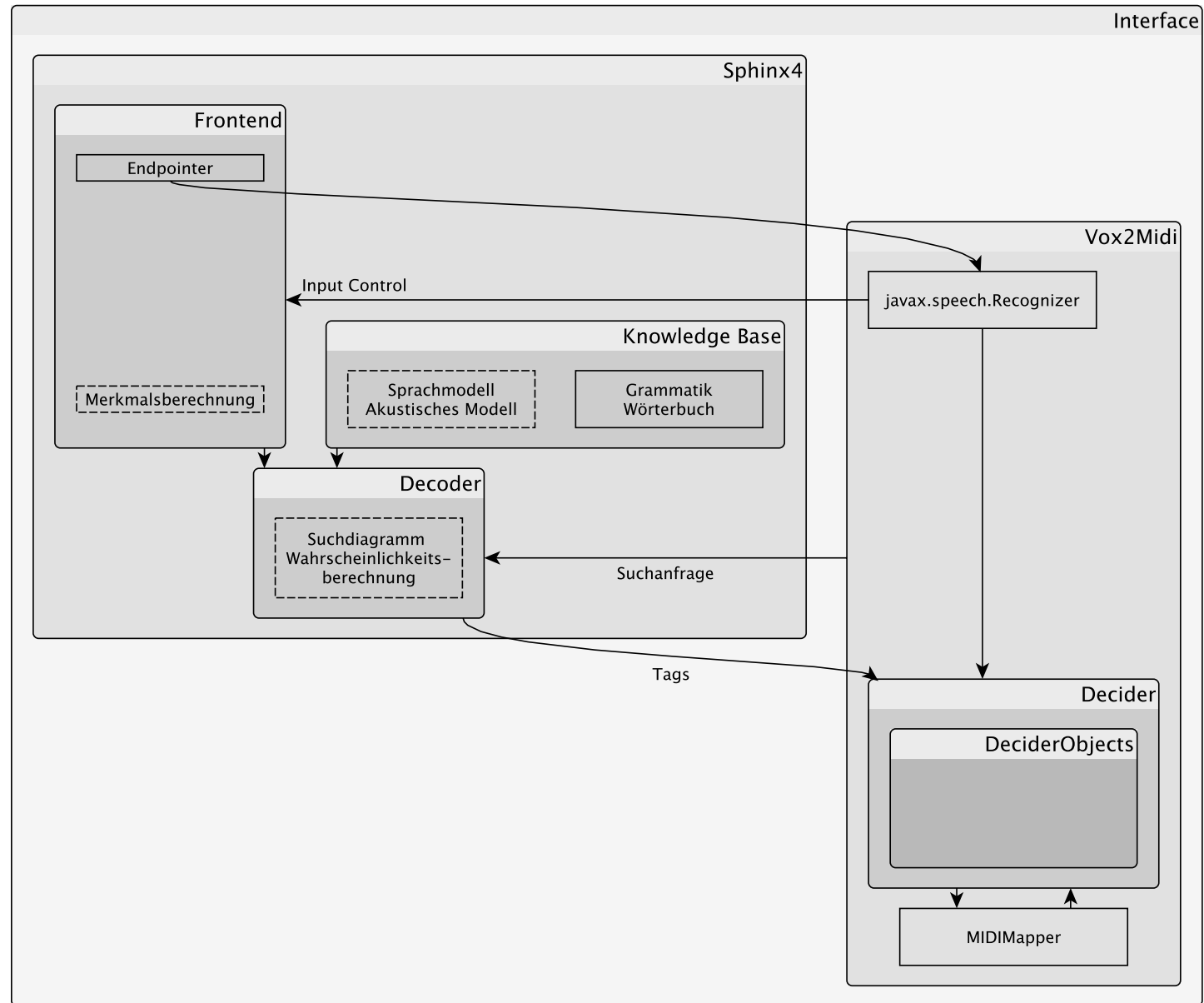
- Steuerdaten
- Sysex
- MMC

## **Aufbau des Systems**

### **Linux System**

- Sphinx 4**
- Vox2Midi**
- Decider**

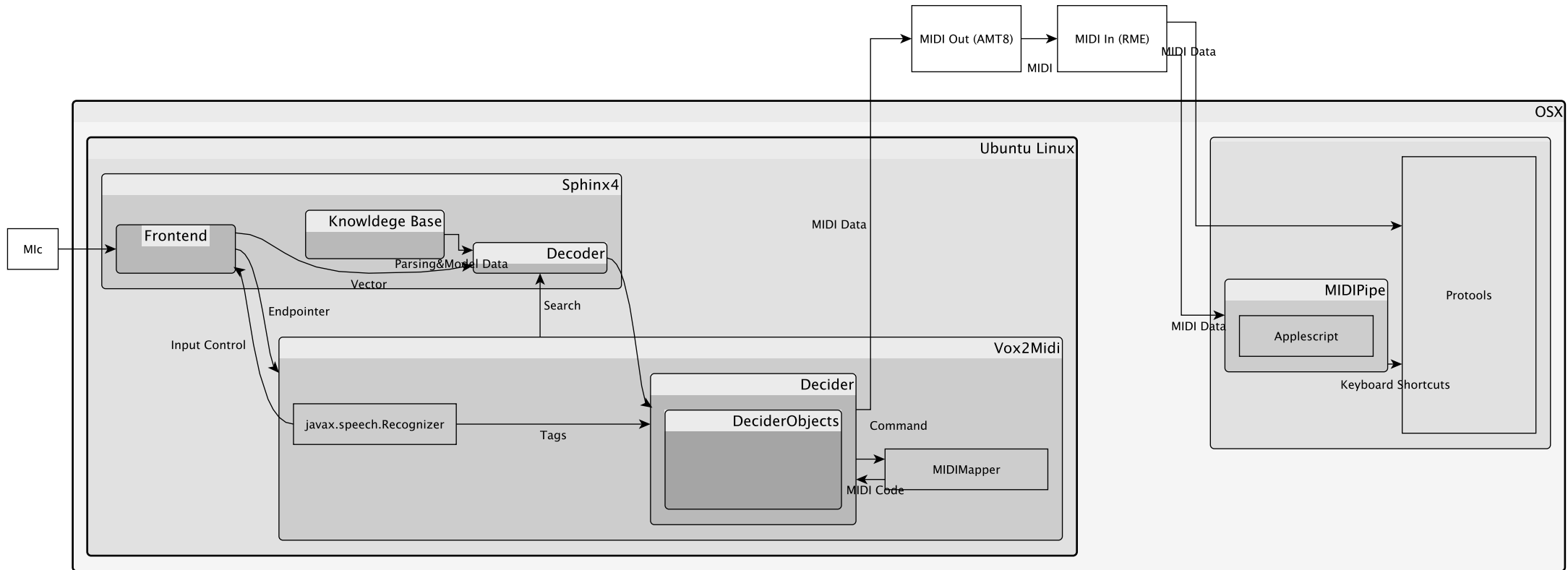
# Aufbau des Systems



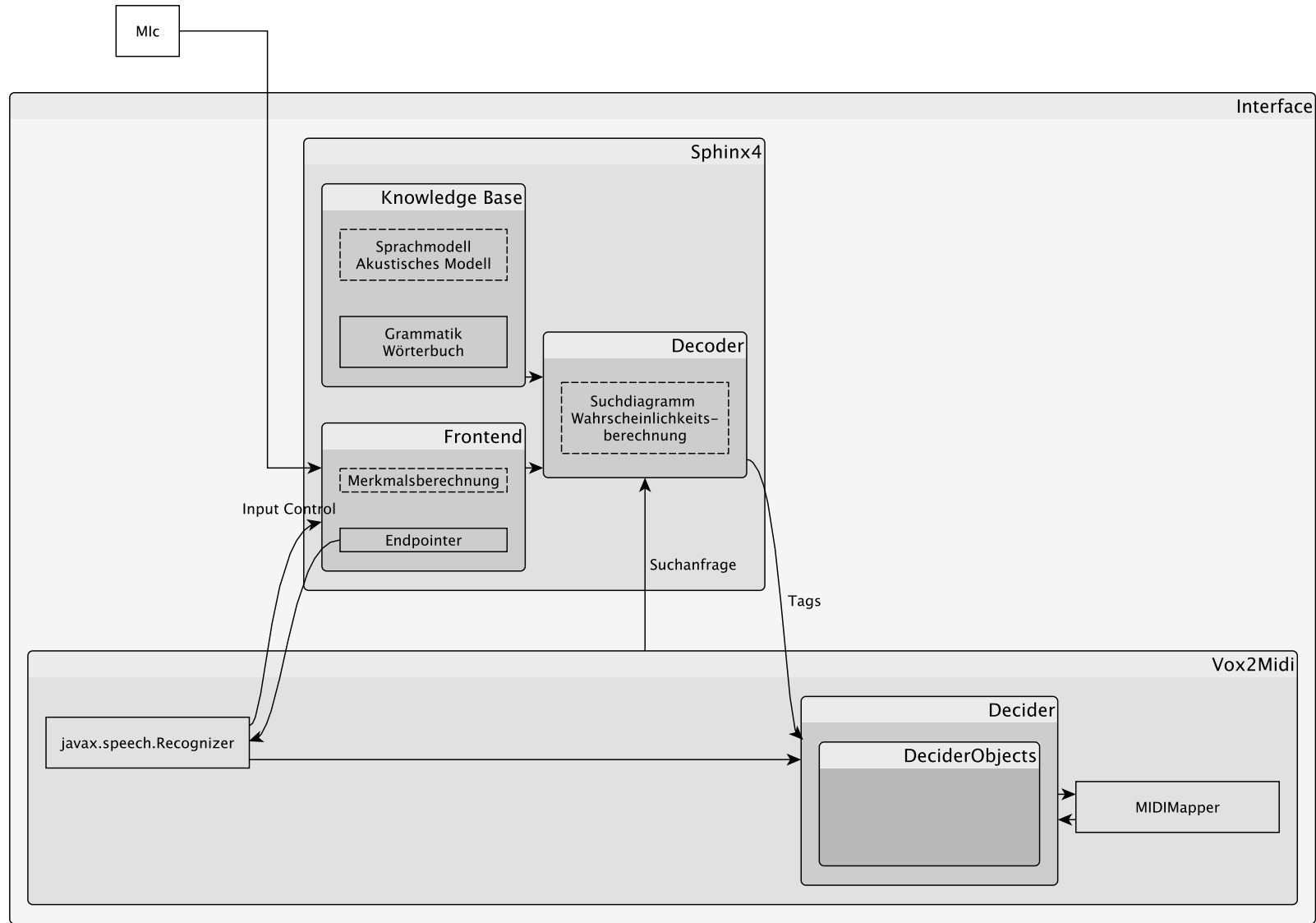
## **Aufbau des Systems**

**Komplettsystem unter OSX  
Linux System in VmWare  
Protools  
Tool: MidiPipe**

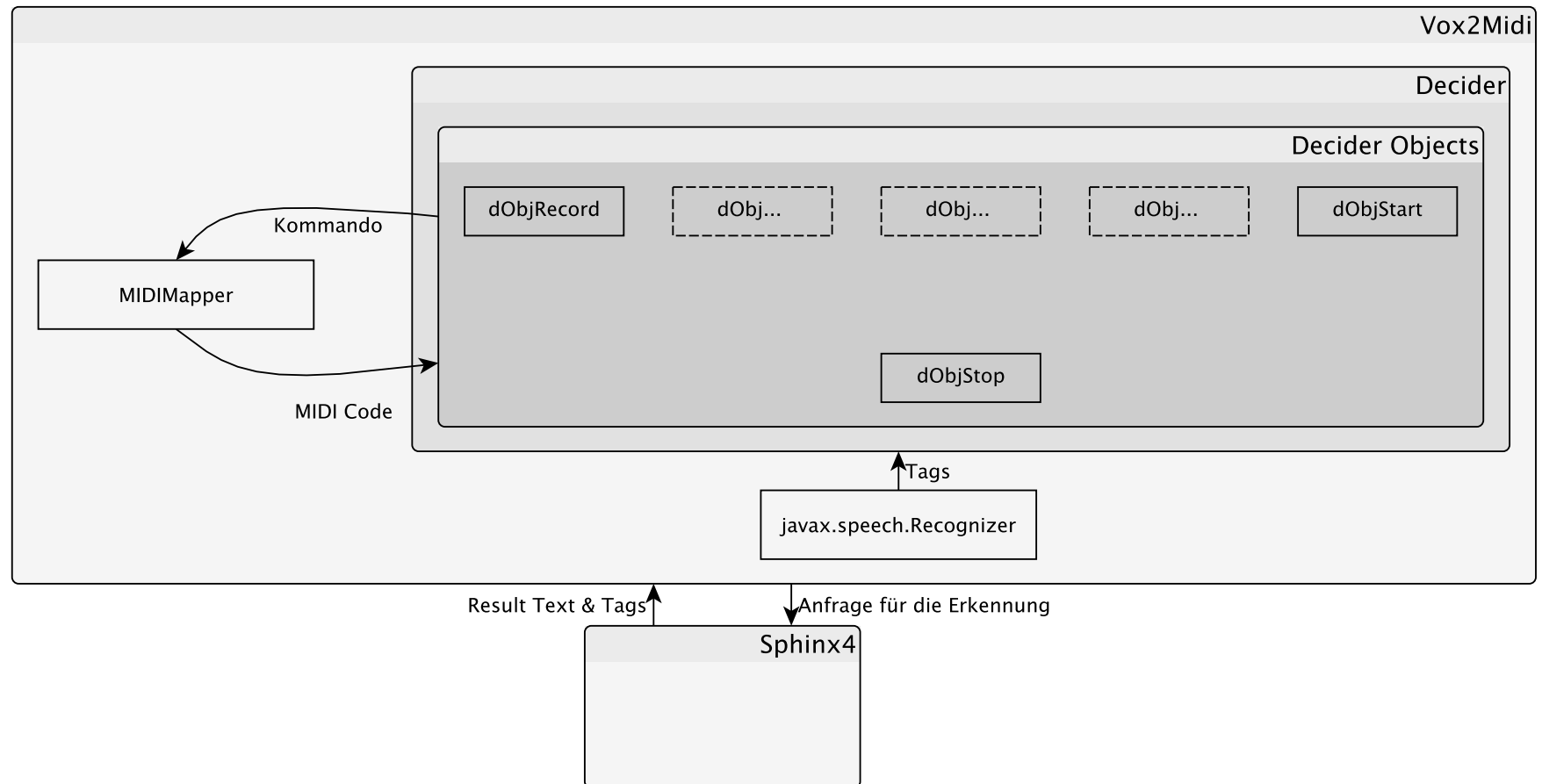
# Aufbau des Systems



# Funktionsweise - Schnittstelle



# Funktionsweise - Decider



# Optimierung

## Grammatik:

```
#JSGF V1.0;  
grammar org.triggerpt.vox2midi;  
/**  
/* simple grammar file example  
*/  
  
public <basiccmd> = <command> [<object>];  
<command> = /10/ start {START} ||5/ (stop | halt) {STOP};  
<object> = [the | a] (song | session) {SONG};
```

# Optimierung

## Wörterbuch

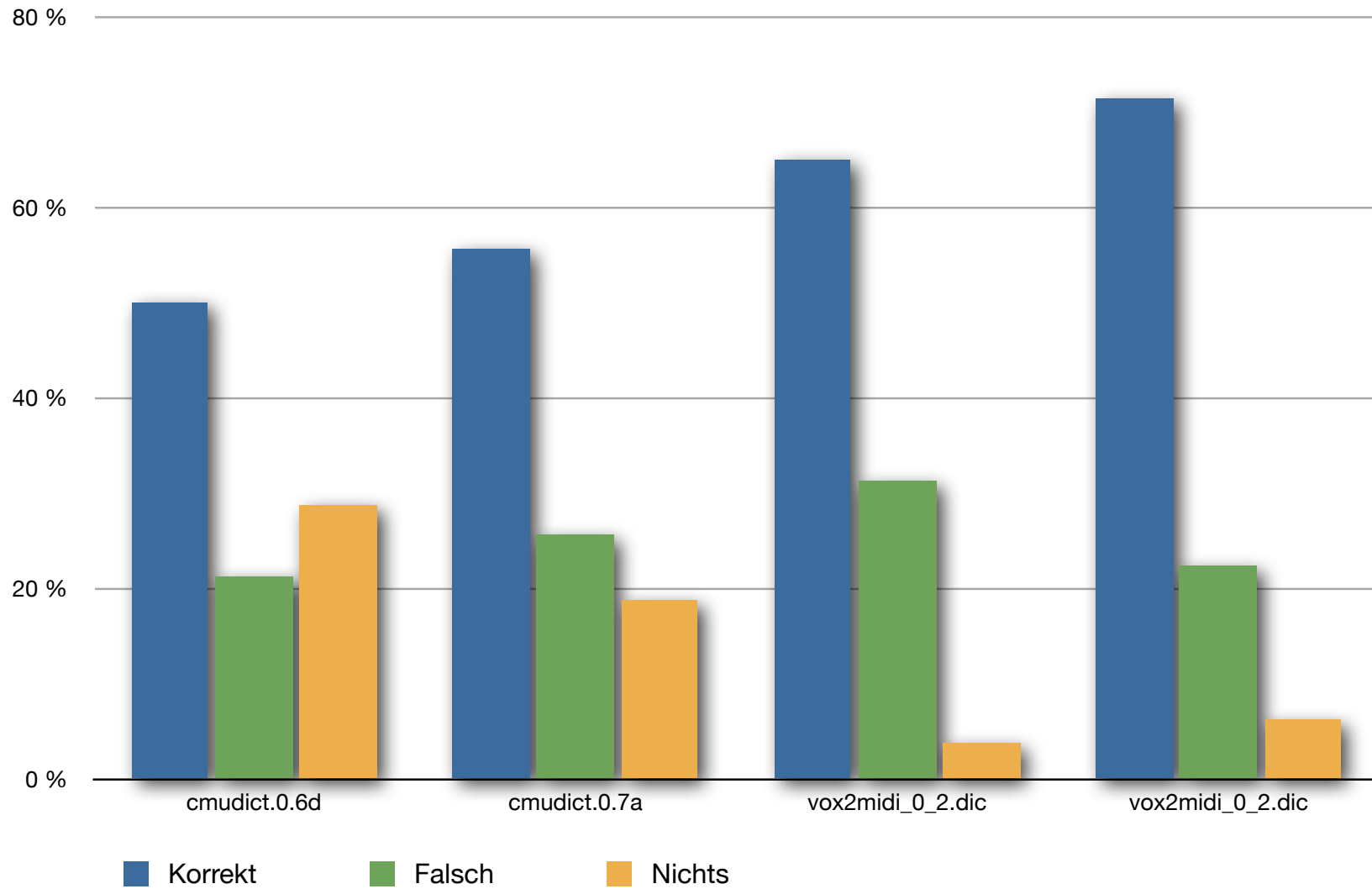
Tool: <http://www.speech.cs.cmu.edu/tools/lmtool-new.html>

A AH  
A(2) EY  
ARM AA R M  
ARM(2) AH R M  
AT AE T  
CAN K AE N  
CAN(2) K AH N  
COMPUTER K AH M P Y UW T ER  
COULD K UH D  
EIGHT EY T  
EIGHTEEN EY T IY N  
EIGHTY EY T IY  
ELEVEN IH L EH V AH N  
ELEVEN(2) IY L EH V AH N  
ENABLEEH N EY B AH L  
ENABLE(2) IH N EY B AH L  
EXIT EH G Z IH T  
EXIT(2) EH K S AH T

·  
·  
·

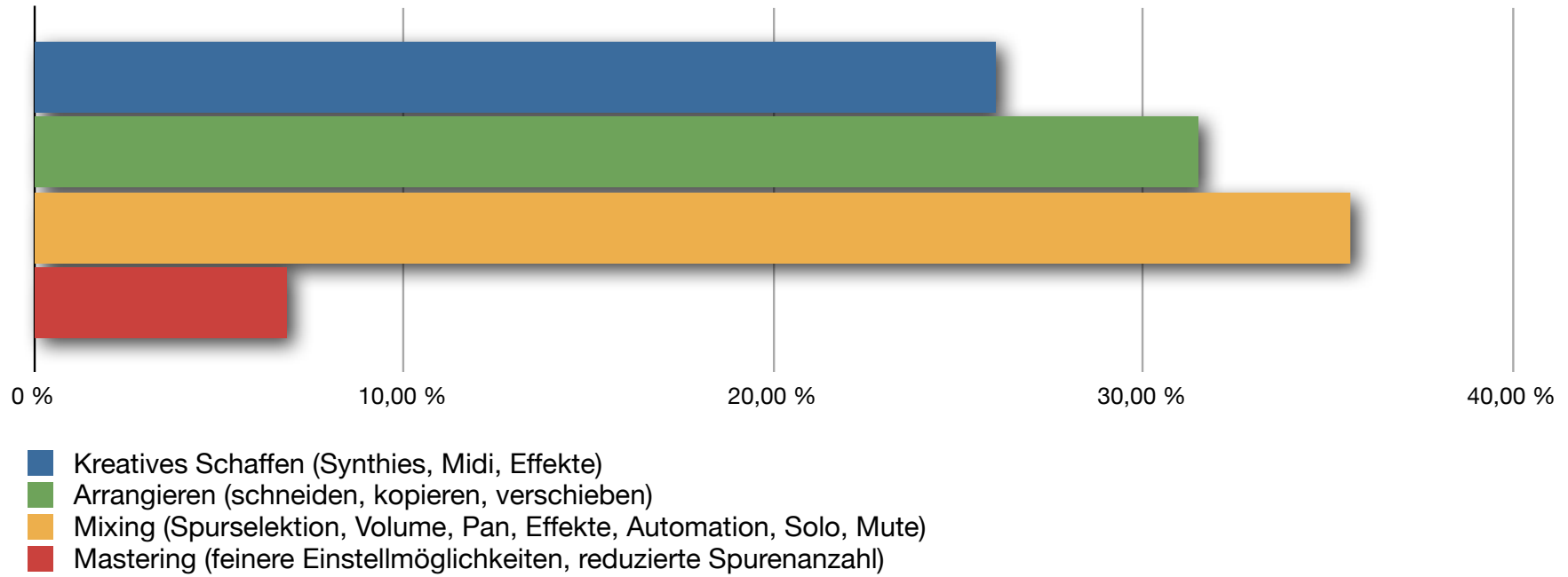
# Optimierung

## Tests der Erkennungsrate



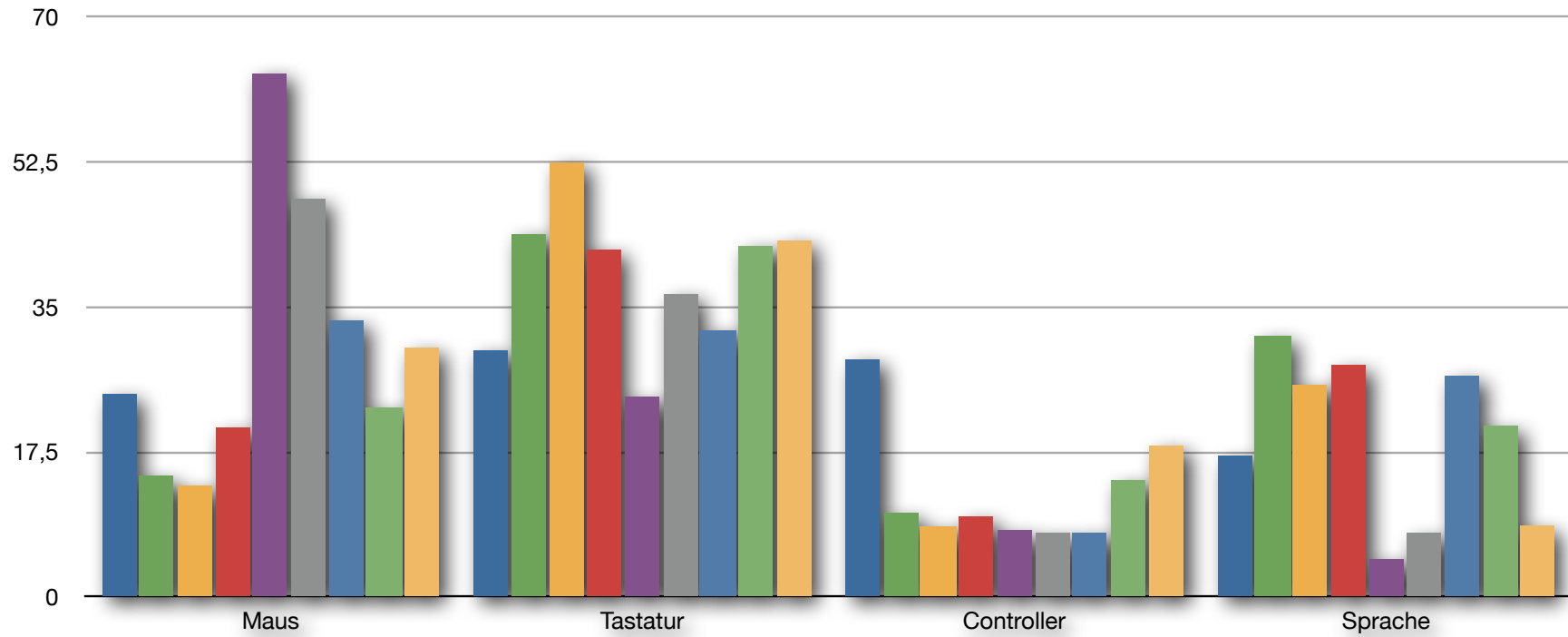
# Umfrage

Frage 5: Für welchen Arbeitsprozess könnten Sie sich eine Sprachsteuerung vorstellen?



# Umfrage

Frage 6: Welche Eingabemethode würden Sie für die jeweiligen Funktionen vorziehen?



- Transport
- Werkzeuganwahl
- Raster, Edit-Modes umschalten
- Anzeigeoptionen (Fenster, Zoom, usw.)
- Auswahl/Markieren
- Editieren, Bewegen, Schneiden
- Menübefehle
- Standardbefehle
- Makrobefehle

**Schlussfolgerungen**

**&**

**Ausblick**

## Links

Skript: [http://www.triggerpt.org/2012/02/ba\\_skript/](http://www.triggerpt.org/2012/02/ba_skript/)

Sphinx 4: <http://cmusphinx.sourceforge.net/sphinx4/>

Speech at the Carnegie Mellon University: <http://www.speech.cs.cmu.edu/>

Java: <http://www.oracle.com/technetwork/java/index.html>

MidiPipe: <http://web.mac.com/nicowald/SubtleSoft/MidiPipe.html>

Ubuntu Linux: <http://www.ubuntu.com>

## Kontakt

Christian Neyens

Tel: 0176 57 09 81 10

Email: [cneyens@triggerpt.org](mailto:cneyens@triggerpt.org)