

# EXPLORATIVE UNTERSUCHUNG VON MULTI-INTENTS IN SPRACHDIALOGSYSTEMEN

*Jakob Landesberger, Daniel Kornmüller, Ute Ehrlich*

*User Interaction, Daimler AG*

*{jakob.landesberger, daniel.d.kornmueller, ute.ehrlich}@daimler.com*

**Kurzfassung:** Mehrere Absichten in einer Äußerung zu formulieren, kann einen Dialog effektiver und schneller machen. Daher wurden im Rahmen einer Internetstudie mit Hilfe von grafischen Anweisungen spontane Nutzeräußerungen gesammelt. Die Aufgaben bewegten sich alle im Bereich des automatisierten Fahrens und verlangten von den Probanden eine variierende Anzahl an Absichten in einer Äußerung zu formulieren. Die transkribierten Nutzeräußerungen wurden hinsichtlich Anzahl der Sätze, syntaktischer Strukturen und Natürlichsprachlichkeit analysiert. Die Probanden wählten unabhängig von der geforderten Anzahl an Absichten in einer Äußerung größtenteils einen von Imperativformen geprägten natürlichsprachlichen Stil. Die Analyse der Äußerungen mit mehr als einer Intention offenbarte Hinweise, dass nicht nur der Dialogeinstieg als Kontext für Multi-Intent-Äußerungen relevant sein kann.

## 1 Einleitung

Bei Sprachdialogsystemen ist ein entscheidender Schritt für den Erfolg der Interaktion das Verstehen der Absicht eines Nutzers. Dieser Schritt wird umso komplizierter, je mehr Absichten ein Nutzer formuliert. Multi-Intent wird dabei als Begriff für mehrere unterschiedliche Absichten, die in einem einzigen natürlichsprachlichen Satz geäußert werden, verwendet. Die Möglichkeit, mehrere Absichten auszudrücken, verspricht analog zur zwischenmenschlichen Kommunikation diverse Vorteile. Denn in einer typischen zwischenmenschlichen Interaktion werden häufig mehrere Absichten in einem Satz geäußert. Die menschliche Fähigkeit, diese Absichten einzuordnen und darauf zu reagieren, erlaubt einen effektiveren und schnelleren Dialog [1].

Der Schwerpunkt der aktuellen Multi-Intent-Forschung liegt dabei meist in der Erkennung und Differenzierung von einzelnen gegenüber mehreren Absichten in einer Äußerung. Kim et al. stellen beispielsweise ein mögliches zweistufiges Verfahren zur Erkennung und Separierung von Multi-Intents vor. Allerdings wird das Verfahren nur mit Hilfe von kombinierten Single-Intents trainiert [2]. Xu et al. nutzen Gemeinsamkeiten zwischen den verschiedenen Absichten in einer Äußerungen um diese besser verstehen zu können [3]. Beaver et al. stellt einen annotierten Korpus der auch Multi-Intents enthält vor [4].

Diese Forschungsarbeiten konzentrieren sich ausschließlich auf den Dialogeinstieg und gehen wenig auf das Nutzerverhalten und die Variationen in der Struktur der gesprochenen Äußerungen ein.

Als einen ersten Schritt in diese Richtung wird hier eine explorative Untersuchung präsentiert, die verschiedene Arten der Formulierung von Multi-Intents aufzeigt und mit einzelnen Absichten in Äußerungen, auch Single-Intents genannt, vergleicht.

## 2 Versuchsaufbau

In einer Internetstudie wurde untersucht wie Probanden Anfragen an ein fiktives System, genannt PISA, formulieren. Das Ziel dabei ist, verschiedene Sprachstile zu identifizieren und syntaktische Unterschiede in Sätzen mit nur einer gegenüber mehreren Absichten herauszuarbeiten.

Die Probanden lösten unbeobachtet am eigenen Endgerät über ein Web-Interface sieben verschiedene Aufgaben. Tabelle 1 gibt einen Überblick über die Aufgabeninhalte und die Anzahl der geforderten Absichten. In fünf Dialogen wurde eine einzige Absicht gefordert, in den zwei übrigen Fragestellungen mussten die Probanden zwei Absichten in eine Äußerung verpacken. Die Aufgaben zielten alle auf eine Interaktion mit einem automatisiert fahrenden Fahrzeug ab. Dabei bewegten sich die Anweisungen zwischen alltäglichen Szenarien, wie ein Navigationsziel einzugeben, und unbekannteren Aufgaben, wie z.B. das Fahrzeug sprachlich zu einer automatischen Parkplatzsuche aufzufordern.

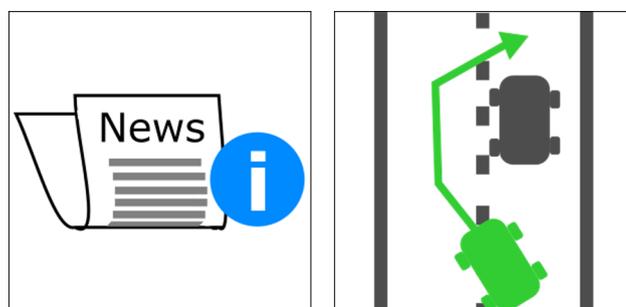
Um einer Beeinflussung im Voraus zu entgehen, wurden die Probanden nur anhand einer klei-

**Tabelle 1** – Übersicht über die Aufgaben und Aufgabentypen.

Nr	Aufgabe	Aufgabentyp
1	Navigationsziel & Keine Musik	Multi-Intent
2	Autonomes Fahren	Single-Intent
3	Aktuelle Nachrichten	Single-Intent
4	Autonomes Überholen	Single-Intent
5	Manuelles Fahren	Single-Intent
6	Zustimmung & Heizung Erhöhen	Mutli-Intent
7	Autonomes Parken	Single-Intent

ner Vorgeschichte und grafischen Aufgabenstellungen instruiert. Exemplarisch zeigt Abbildung 1 die Anweisung zu Aufgabe drei und vier. Damit die Bilder leichter interpretierbar sind, wurden die einzelnen Fragestellungen in eine Geschichte verpackt, bei der der Proband als Protagonist fungiert und den Weg von der Haustür bis zur Arbeitsstelle mit einem hochautomatisierten Fahrzeug erlebt.

Der Großteil der Dialoge (sechs aus sieben) waren nutzerinitiiert. Dadurch wurde eine nut-



**Abbildung 1** – Grafische Aufgabenstellung zu Aufgabe 3 und 4.

zerseitige Beeinflussung durch den Sprachstil der Systemausgaben verhindert. Denn mehrere Untersuchungen belegen, dass Nutzer von Sprachdialogsystemen sich sehr stark an lexikalischen und syntaktischen Entscheidungen von Systemausgaben orientieren [5] [6].

### 3 Datenanalyse

Die gesammelten Daten wurden von Hand hinsichtlich Anzahl der Sätze und Funktionen der Sätze bzw. der Verben analysiert.

Tabelle 2 zeigt die verschiedenen Klassifikationen für Äußerungen mit einer oder mehreren Absichten. Treten mehr als eine Intention in einem Ausdruck auf, werden die verschiedenen Sätze separat markiert. Dabei können syntaktische Abhängigkeiten auftreten, die bei isoliert betrachteten Sätzen zu falschen Schlussfolgerungen führen können.

Äußerung 1: „Ich möchte zur Arbeit fahren“ [...]  
Äußerung 2: [...] „dabei aber keine Musik hören.“

Beispielsweise wird die elliptische Äußerung 2 in Kombination mit dem ersten Teilausdruck (Äußerung 1) ebenfalls als Modalverbssatz klassifiziert, da das fehlende flektierte Verb im ersten Satz auftritt und eine Wiederholung sprachlich nicht erforderlich ist.

**Tabelle 2** – Übersicht über die Klassifikationen mit einem Beispiel aus den entsprechenden Aufgaben.

Klassifikation	Beispiel	Aufgabe
Verbloser Satz, Verbloser Satz	Einmal zur Arbeit. Ton aus.	1
Imperativsatz, Modalverbssatz	Bitte bringe mich zur Arbeit. Ich möchte auf dem Weg dahin keine Musik hören.	1
Infinitivsatz	Bitte autonom weiterfahren.	2
Imperativsatz	Zeige mir die aktuellen Nachrichten.	3
Interrogativsatz	Welche Neuigkeiten gibt es heute?	3
Indikativsatz	Den Traktor kannst du überholen.	4
Modalverbssatz	Ich möchte wieder selbst fahren.	5
Imperativsatz, Imperativsatz	Plane einen Zwischenstopp bei einer Tankstelle ein und dreh die Klimaanlage auf 25 Grad.	6
Verbloser Satz	Autonomes Parken.	7

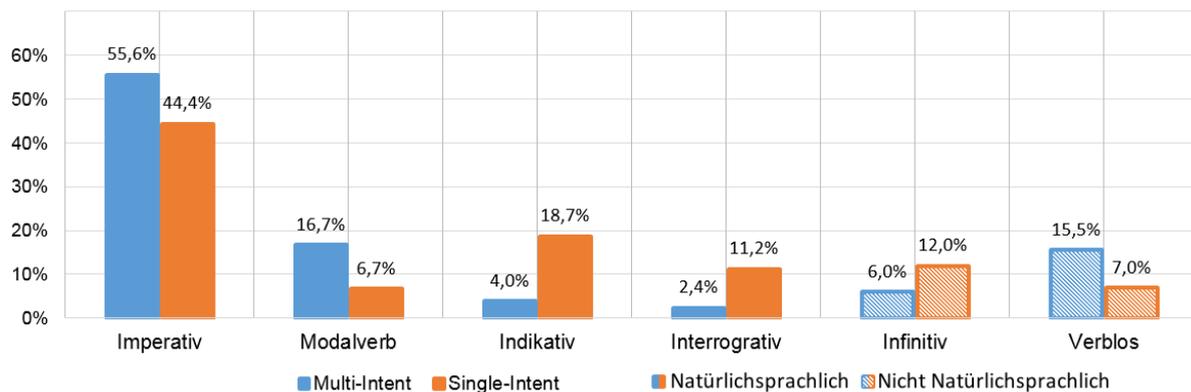
In der sprachlichen Interaktion mit einer Maschine wird häufig zwischen verschiedenen Sprachstilen von Nutzern unterschieden. Berg et al. und Hofmann et al. differenzieren beispielsweise zwischen einem an die zwischenmenschliche Kommunikation angelehnten natürlichen Stil und einen zur Interaktion mit einer Maschine angepasste, kommando- oder schlüsselwortbasierte Sprechweise [7] [8]. Zur Unterscheidung kann hier in beiden Fällen das Vorhandensein bzw. die Flexion des Verbes in den jeweiligen Sätzen herangezogen werden. Braunger et al. unterstützen eine syntaktische Analyse entsprechend der Satzarten und konnten u.a. zeigen, dass frei mit einem System sprechende Nutzer zur Verwendung von Imperativformen neigen [9]. Angelehnt an Berg et al. und Hofmann et al. wurden die gefunden syntaktischen Klassen für eine Untersuchung der Natürlichsprachlichkeit in zwei Gruppen unterteilt. Die erste Gruppe vereint alle Satztypen mit flektiertem Verb und bildet die Gruppe der natürlichsprachlichen Äußerungen. Bei fehlendem oder im Infinitiv auftretendem Verb wird die Äußerung als nicht natürlichsprachlich gewertet.

### 4 Resultate

An der Internetstudie nahmen insgesamt 90 Probanden (36 f / 54 m) mit einem Altersdurchschnitt von 30,01 Jahren ( $\sigma = 12,42$ ) teil. Deren Erfahrung im Umgang mit Sprachdialogsystemen ist gering bis mittel (fünfstufige Likert-Skala;  $\bar{X} = 2,85$ ;  $\sigma = 1,08$ ). Die Probanden nutzen

ein Sprachdialogsystem durchschnittlich ein bis dreimal im Monat (sechsstufige Likert-Skala;  $\bar{X} = 2,00$ ;  $\sigma = 1,25$ ).

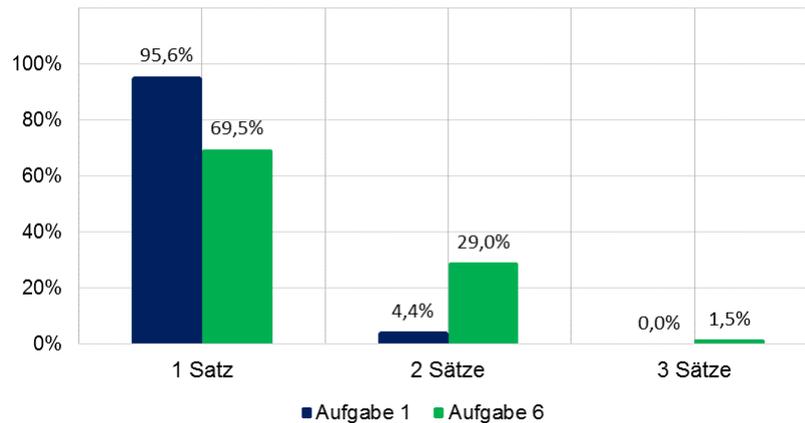
Abbildung 2 zeigt die Verteilung der syntaktischen Klassifizierung einzelner Probandenaussagen unterteilt in Multi- und Single-Intents. Aus der Grafik geht hervor, dass die meisten Teilnehmer sowohl bei Multi-Intents als auch bei Single-Intents einen Imperativsatz bevorzugen. Bei Multi-Intents liegt der Anteil bei 55,6 % und bei Single-Intents bei 44,4%. Bei nur einer Absicht in einer Äußerung sind die weiteren Tendenzen Sätze im Indikativ (18,7 %) oder mit Infinitiv (12,0 %). Im Gegensatz dazu wurden bei Multi-Intents Modalverben (16,7 %) oder verblose Sätze (15,5 %) bevorzugt. Interrogativsätze treten bei den analysierten Daten fast ausschließlich bei Single-Intents auf (11,2 %; 2,4 %).



**Abbildung 2** – Verteilung der syntaktischen Klassifizierung für Single- und Multi-Intents.

Abbildung 3 beschreibt die Anzahl der Sätze, die für eine Intention bei einer Multi-Intent Äußerung verwendet wurden. Dabei werden die zwei Aufgaben, bei denen mehr als eine Absicht in einer Aussage gefordert war, gegenüber gestellt. Die beiden Aufgaben unterscheiden sich hinsichtlich des Dialogkontextes. Während die erste Fragestellung auf einen nutzerinitiierten Einstieg in den Dialog abzielt, wird bei Aufgabe sechs das aktuelle Thema des andauernden Diskurses mit Hilfe eines Multi-Intents erweitert. Bei beiden Aufgaben tendieren die Probanden hauptsächlich dazu, einen Satz pro Intension zu verwenden (95,6 %; 69,5 %). Bei Aufgabe sechs nutzen allerdings auch 29,0 % zwei Sätze, um eine Absicht zu formulieren. Der zusätzliche Satz wurde hauptsächlich genutzt, um die aus der Zustimmung resultierende Absicht zu spezifizieren (Äußerung 3) oder die gewünschte Handlung zu begründen (Äußerung 4).

- Äußerung 3: „Gute Idee! Erwinnere mich später daran.“ [...]  
 Äußerung 4: [...] „Bitte dreh die Heizung auf. Mir ist kalt.“



**Abbildung 3** – Übersicht über die Anzahl der verwendeten Sätze für eine Absicht bei Multi-Intents.

Bei dem Vergleich der Natürlichsprachlichkeit von Äußerungen mit nur einer Absicht gegenüber solchen mit mehreren Intensionen zeigte sich kein Unterschied ( $p = 0,7943$ ). Bei Single-Intent Aufgaben zeichnete sich bei 81,0 % der Äußerungen ein entsprechend in Kapitel 3 definierter natürlicher Sprachstil ab. Bei Fragestellungen mit Multi-Intents entsprachen 78,7 % natürlichsprachlichen Charakteristiken.

## 5 Diskussion

Die beobachtete ähnliche Verteilung von natürlichsprachlichen Äußerungen lässt darauf schließen, dass es Probanden nicht schwerer fällt, mehrere Absichten in einer Äußerung zu formulieren, als die Intentionen isoliert auszusprechen. Das unterstreicht das Potential durch Verwendung von Multi-Intents einen effektiveren und schnelleren Dialog zu ermöglichen.

Die in Abbildung 2 vorgestellte Verteilung der syntaktischen Klassifizierung zeigt, dass sich Sätze bei Äußerungen mit mehreren Absichten anders zusammensetzen als bei isolierten Single-Intents. Es ist allerdings nicht auszuschließen, dass das Auftreten der hier beobachteten Varianz durch die geringe Variation und schlechte Vergleichbarkeit der Aufgabentypen begünstigt wird. Nichts desto trotz weisen die Ergebnisse auf nicht zu vernachlässigende syntaktische Unterschiede hin. Das legt den Schluss nahe, dass sich durch das Kombinieren von Single-Intents keine validen Multi-Intents erzeugen lassen, sondern separate Korpora anhand von spontan gesprochenen Nutzenäußerungen erstellt werden sollten.

Ähnlich verhält es sich mit der beobachteten unterschiedlichen Anzahl der analysierten Sätze bei Aufgaben mit Multi-Intents. Diese Varianz kann an Hand der hier erhobenen Daten nicht mit definitiver Sicherheit ausschließlich auf den veränderten Dialogkontext zurückgeführt werden. Trotzdem deutet die Untersuchung auf die Wichtigkeit des Dialogkontextes hin und offenbart weitere Möglichkeiten, wie Multi-Intents effektiv in einem Dialog eingesetzt werden können.

## 6 Fazit

Im Rahmen dieser Studie konnten Indizien festgestellt werden, dass Probanden bei der Äußerung von mehrfachen Absichten in einem Satz genauso wie bei einzelnen Intentionen größtenteils einen natürlichen, von Imperativformen geprägten Sprachstil wählen. Zusätzlich konnten Hinweise gefunden werden, dass der Dialogkontext bei Multi-Intents eine Rolle spielt, und der Fokus zukünftiger Forschung nicht nur auf dem Dialogeinstieg, sondern auch auf der Fortführung bzw. Erweiterung von Dialogen, gelegt werden sollte. Neben dem Dialogkontext müssen die aus der Verwendung von Multi-Intents resultierenden, für den Nutzer nötigen, Adaptionen

in der Dialogführung und von Feedbackstrategien untersucht werden. Das alles ist Ziel von zukünftigen Forschungen.

## Literatur

- [1] LEMON, O., A. GRUENSTEIN, A. BATTLE, und S. PETERS: *Multi-tasking and collaborative activities in dialogue systems*. In *Proceedings of the 3rd SIGdial workshop on Discourse and dialogue-Volume 2*, S. 113–124. Association for Computational Linguistics, 2002.
- [2] KIM, B., S. RYU, und G. G. LEE: *Two-stage multi-intent detection for spoken language understanding*. *Multimedia Tools and Applications*, 76(9), S. 11377–11390, 2017.
- [3] XU, P. und R. SARIKAYA: *Exploiting shared information for multi-intent natural language sentence classification*. In *INTERSPEECH*, S. 3785–3789. 2013.
- [4] BEAVER, I., C. FREEMAN, und A. MUEEN: *An annotated corpus of relational strategies in customer service*. *arXiv preprint arXiv:1708.05449*, 2017.
- [5] BRANIGAN, H. P., M. J. PICKERING, J. PEARSON, und J. F. MCLEAN: *Linguistic alignment between people and computers*. *Journal of Pragmatics*, 42(9), S. 2355–2368, 2010.
- [6] ZOLTAN-FORD, E.: *How to get people to say and type what computers can understand*. *International Journal of Man-Machine Studies*, 34(4), S. 527–547, 1991.
- [7] HOFMANN, H., U. EHRLICH, A. BERTON, und W. MINKER: *Speech interaction with the internet—a user study*. In *Intelligent Environments (IE), 2012 8th International Conference on*, S. 323–326. IEEE, 2012.
- [8] BERG, M., P. GRÖBER, und M. WEICHT: *User study: Talking to computers*. In *3rd Workshop on inclusive eLearning, London (UK)*, Bd. 9. 2010.
- [9] BRAUNGER, P., H. HOFMANN, S. WERNER, und M. SCHMIDT: *A comparative analysis of crowdsourced natural language corpora for spoken dialog systems*. In *LREC*. 2016.