

DIN EN ISO 9241-154

ICS 13.180; 35.240.99

Einsprüche bis 2011-06-04

Entwurf

**Ergonomie der Mensch-System-Interaktion –
Teil 154: Dialogführung mittels Sprachdialogsystemen
(ISO/DIS 9241-154:2011);
Deutsche Fassung prEN ISO 9241-154:2011**

Ergonomics of human-system interaction –
Part 154: Interactive voice response (IVR) applications (ISO/DIS 9241-154:2011);
German version prEN ISO 9241-154:2011

Anwendungswarnvermerk

Dieser Norm-Entwurf mit Erscheinungsdatum 2011-03-28 wird der Öffentlichkeit zur Prüfung und Stellungnahme vorgelegt.

Weil die beabsichtigte Norm von der vorliegenden Fassung abweichen kann, ist die Anwendung dieses Entwurfes besonders zu vereinbaren.

Stellungnahmen werden erbeten

- vorzugsweise als Datei per E-Mail an naerg@din.de in Form einer Tabelle. Die Vorlage dieser Tabelle kann im Internet unter www.din.de/stellungnahme oder für Stellungnahmen zu Norm-Entwürfen der DKE unter www.dke.de/stellungnahme abgerufen werden;
- oder online im Norm-Entwurfs-Portal des DIN unter www.entwuerfe.din.de, sofern dort wiedergegeben;
- oder in Papierform an den Normenausschuss Ergonomie (NAErg) im DIN, 10772 Berlin (Hausanschrift: Burggrafenstr. 6, 10787 Berlin).

Die Empfänger dieses Norm-Entwurfs werden gebeten, mit ihren Kommentaren jegliche relevanten Patentrechte, die sie kennen, mitzuteilen und unterstützende Dokumentationen zur Verfügung zu stellen.

Gesamtumfang 48 Seiten

Normenausschuss Ergonomie (NAErg) im DIN
Normenausschuss Informationstechnik und Anwendungen (NIA) im DIN

Nationales Vorwort

Dieses Dokument (prEN ISO 9241-154:2011) wurde vom Technischen Komitee ISO/TC 159, Unterkomitee SC 4 „Ergonomics of human-system interaction“, Arbeitsgruppe WG 5 „Software ergonomics and human-computer dialogues“ in Zusammenarbeit mit dem Technischen Komitee CEN/TC 122 „Ergonomie“ erarbeitet, dessen Sekretariat vom DIN (Deutschland) gehalten wird. Das zuständige deutsche Gremium ist der NA 023-00-04-05 GAK „Gemeinschaftsarbeitskreis NAErg/NIA; Benutzungsschnittstellen“ im Normenausschuss Ergonomie (NAErg).

Das Ausgabedatum des Europäischen Norm-Entwurfs stand zum Zeitpunkt der Veröffentlichung dieses DIN-EN-ISO-Entwurfs noch nicht fest; der prEN ISO wird jedoch vom CMC unter der angegebenen prEN-Nummer demnächst zur CEN-Umfrage verteilt. Um der deutschen Öffentlichkeit eine möglichst lange Einspruchsfrist einzuräumen, wurde dieser Norm-Entwurf bereits vorab veröffentlicht.

ISO 9241 wurde ursprünglich als eine siebzehnteilige Norm zu ergonomischen Anforderungen an Büro-tätigkeiten mit Bildschirmgeräten erarbeitet. Als Teil des Normenüberarbeitungsprozesses wurde eine größere Neustrukturierung der ISO 9241 vereinbart, um den Anwendungsbereich zu erweitern, weitere relevante Normen einzubeziehen und die Nutzbarkeit zu verstärken. Der Titel der überarbeiteten ISO 9241 „Ergonomie der Mensch-System-Interaktion“ spiegelt diese Änderungen wieder und richtet die Norm an dem Haupttitel und dem Anwendungsbereich des SC 4 aus. Die überarbeitete mehrteilige Norm ist als eine Reihe von Normen aufgebaut, die in „Hundertern“ beziffert ist; z. B. behandelt die Normenreihe ISO 9241-100 die Software-Schnittstellen, die Normenreihe ISO 9241-200 das auf den Menschen gerichtete Design, die Normenreihe ISO 9241-300 behandelt Sichtanzeigen, die Normenreihe ISO 9241-400 behandelt physikalische Eingabegeräte usw.

Ergonomie der Mensch-System-Interaktion — Teil 154: Dialogführung mittels Sprachdialogsystemen (ISO/DIS 9241-154:2011)

*Ergonomics of human-system interaction — Part 154: Interactive voice response (IVR) applications
(ISO/DIS 9241-154:2011)*

ICS:

Deskriptoren:

Dokument-Typ: Europäische Norm
Dokument-Untertyp:
Dokumentstufe: parallele Umfrage
Dokumentsprache: D

STD Version 2.4c

Inhalt

Seite

Vorwort	5
Einleitung.....	7
1 Anwendungsbereich	9
2 Normative Verweisungen.....	9
3 Begriffe	9
4 Übereinstimmung	16
5 Anwendungen zur Anrufaufzeichnung (en: voice messaging).....	17
6 Eingabe von Information.....	17
6.1 Allgemeines	17
6.2 Informieren von Texttelefon-Benutzern über mögliche Eingabearten.....	17
6.3 Keine Vervielfältigung der Eingabe von Information	17
6.4 Verwenden von Systeminformation zur Verbesserung der Leistungsfähigkeit	17
6.5 Ändern von eingegebener Information	18
7 Spracheingabe	18
7.1 Wortschatzwahl für sprachgesteuerte IVR	18
7.2 Phonetisch unterschiedlicher Wortschatz bei sprachgesteuerten IVR-Anwendungen	18
7.3 Zustandspezifische Grammatik.....	18
7.4 Synonyme in Grammatiken	18
7.5 Verwendung dynamischer Grammatiken.....	19
7.6 Enrollment (das Spracherkennungssystem trainieren).....	19
7.6.1 Allgemeines	19
7.6.2 ASR-Training in der Umgebung des Benutzers	19
7.6.3 ASR-Training durch Wiederholen	19
7.7 End-Pointing setzen und Sprecherwechsel.....	19
7.8 Handhabung von Zeitüberschreitungen bei Spracheingabe	20
7.9 Sprach-IVR mit Tonwahl	20
8 Tonwahleingabe.....	20
8.1 Anzeigen der Tonwahlfähigkeit.....	20
8.2 Tastenbelegung für Begrenzungszeichen	20
8.3 Die Rautetaste (#) bei Dateneingaben mit festgelegter Länge.....	20
8.4 Handhabung von Zeitüberschreitungen für Tonwahleingabe unterschiedlicher Länge	20
8.5 Tastenbelegung für positive und negative Antworten.....	21
8.6 Tastenbelegung für die „Hilfe“-Funktion	21
8.7 Minimieren von Anrufer-Tastendrücken	21
9 Ausgabe von Information	21
9.1 Allgemeine Festlegungen für Prompts und Ansagen.....	21
9.1.1 Kürze der Information	21
9.1.2 Gruppieren von Information	21
9.1.3 Betonung und Satzmelodie	21
9.1.4 Sprechgeschwindigkeit.....	22
9.1.5 Einfache linguistische Formen.....	22
9.1.6 Vernünftige Anwendung höflicher Sprache bei Prompts.....	22
9.1.7 Beenden des Anrufs	22
9.1.8 Akustische Qualität von Prompts und Ansagen	22
9.1.9 Akustisch einheitliche Lautstärke von Prompts und Ansagen	23
9.1.10 Anruferwahrnehmung des IVR als Computer.....	23
9.1.11 Persona.....	23

9.1.12	Wiederholung von Prompts und Ansagen.....	23
9.1.13	Wiederholung von Teilen langer Prompts und Ansagen	23
9.1.14	Prompts und Ansagen für Benutzer von Texttelefonen.....	24
9.1.15	Einheitlichkeit der Sprache	24
9.1.16	Sprache von Texttelefonmeldungen (TTY)	24
9.2	Konstruktion von Prompts und Ansagen	24
9.2.1	Sortieren von Einträgen in Optionslisten	24
9.2.2	Sortieren von Tastenbelegungen für Optionslisten	24
9.2.3	Sortieren von Zielen und Aktionen in Prompts.....	25
9.2.4	Betonung und Satzmelodie	25
9.2.5	Betonung und Satzmelodie von verknüpften Prompts	25
9.2.6	Betonung und Satzmelodie, um Sprecherwechsel bei sprachgesteuerten Anwendungen zu erleichtern	25
9.2.7	Einleitende Ausdrucksweise für Sprachoptionslisten	25
9.2.8	Einschränkende Spracheingabe.....	25
9.2.9	Verwendung von Hinweiswörtern und Orientierungspunkten	26
9.2.10	Eindeutige Prompts.....	26
9.2.11	Unvollständige Gruppe von Optionen.....	26
9.2.12	Zusätzliche Optionen	26
9.2.13	Vermeiden expliziter Anweisungen über die Sprechweise.....	26
9.2.14	In Prompts verwendete Grammatik und Begriffe.....	26
9.2.15	Irrelevante oder nicht verfügbare Optionen	28
9.2.16	Tastenbelegung für nicht verfügbare/irrelevante Optionen.....	28
9.2.17	Einheitlichkeit bei Tastenbelegungen	28
9.2.18	Einheitlichkeit zwischen Sprach- und Tonwahloptionen	28
10	Navigation	29
10.1	Weiterschaltfunktion	29
10.2	Dial-Through-Funktion und Talk-Through-Funktion	29
10.3	Dial-Ahead-Funktion und Talk-Ahead-Funktion.....	29
10.4	Globale Kommandos.....	29
10.5	Barge-In-Funktion.....	29
10.6	Einschränken der Barge-In-Funktion	29
11	Hilfe	30
11.1	Kontextabhängige Hilfe	30
11.2	Vom Benutzer wählbare Hilfe	30
11.3	Bezugnahme auf Hilfe im ersten Prompt	30
11.4	Hilfebefehle	30
11.5	Vom System eingeleitete Hilfe	30
12	Erreichbarkeit eines Mitarbeiters.....	31
12.1	Zugang zu einem Mitarbeiter.....	31
12.2	Gleiche Voraussetzung zur Erreichbarkeit eines Mitarbeiters	31
12.3	Hinweis auf Erreichbarkeit eines Mitarbeiters	31
12.4	Einleiten der Verbindung mit einem Mitarbeiter	31
12.5	Verzögerungen beim Erreichen eines Mitarbeiters	32
12.6	Unterstützung durch einen Mitarbeiter	32
12.7	Keine Hilfe eines Mitarbeiters verfügbar.....	32
13	Rückmeldung	32
13.1	Allgemeines	32
13.2	Rückmeldung auf Benutzereingabe	32
13.3	Rückmeldung, die der Auswahl nicht verfügbarer Optionen folgt.....	33
13.4	Reaktionszeit des Systems	33
13.5	Geeigneter Kontext für Rückmeldungsmittelungen	33
13.6	Orientierungspunkte	33
13.7	Bestätigungsdialoge	34
13.7.1	Allgemeines	34
13.7.2	Keine exzessive Bestätigung	34
13.7.3	Bestätigung einfacher Wörter und Sätze	34

13.7.4	Explizite Eingabebestätigung zur Vermeidung nachteiliger Systemaktionen	34
13.7.5	Bestätigung der Spracheingabe in Abhängigkeit von Konfidenzwerten.....	34
13.7.6	Kombinieren von Bestätigungen	35
13.7.7	Einleitungen für Bestätigungsprompts	35
13.7.8	Wortwahl für Bestätigungsprompts und Ansagen.....	35
13.7.9	Bestätigung von Ziffern- oder Zahlenfolgen	35
13.8	Intelligente Verwendung abgelehnter Bestätigungsfragen.....	35
13.9	Zusätzliche Rückmeldung in Verzögerungssituationen.....	36
13.10	Handhabung ausgedehnter Wartezeiten	36
13.11	Vorschlag, dass der Anrufer wieder anruft.....	36
14	Fehler	36
14.1	Allgemeines	36
14.2	Dienst nicht verfügbar.....	36
14.3	Wiederholungen des Prompts.....	37
14.4	Inhalt von Fehlermeldungen.....	37
14.5	Zwei aufeinanderfolgende Einfügingsfehler bei Tonwahlgesteuerten IVR	37
14.6	Zwei oder mehr aufeinanderfolgende Einfügingsfehler bei sprachgesteuerten IVR	37
14.7	Handhabung von zahlreichen Fehlern.....	37
14.8	Abwertende Wortwahl in Fehlermeldungen.....	37
14.9	Wichtige Mitteilungen.....	38
14.10	Fehlerbehebung	38
14.11	Unterbrechungsmitteilung.....	38
14.12	Unterbrechen infolge Überschreiten der Fehlergrenze	38
14.13	Hinweis auf Hilfe	38
14.14	Wortwahl in Fehlermeldungen	38
14.15	Hinweise über erwünschte Spracheingaben	38
14.16	Beiläufige Fehlerkorrektur	39
14.17	Handhabung von zu viel Spracheingabe	39
14.18	Erwähnung von Tonwahloptionen bei sprachgesteuerten IVR	39
Anhang A (informativ)	Überblick über die Normenreihe ISO 9241	40
Anhang B (informativ)	Die Rolle von Spracherkennungsfehlern bei der Gestaltung von IVR.....	41
B.1	Einleitung — Sprachgesteuerte IVR bringen Erkennungsfehler mit sich.....	41
B.2	Arten von Erkennungsfehlern und ihre Folgen	41
B.2.1	Allgemeines	41
B.2.2	Korrekte Zeitüberschreitung	42
B.2.3	Einfügingsfehler	42
B.2.4	Löschungsfehler	42
B.2.5	Verwechslungsfehler.....	42
B.2.6	Zurückweisungsfehler.....	42
B.2.7	Falsche Erkennung.....	43
B.2.8	Korrekte Zurückweisung.....	43
B.2.9	Häufigkeit der Fehlerarten	43
B.3	Gestaltung der Sprachbenutzungsschnittstelle für die Handhabung von Erkennungsfehlern	43
Literaturhinweise		45

Vorwort

Dieses Dokument (prEN ISO 9241-154:2011) wurde vom Technischen Komitee ISO/TC 159 „Ergonomics“ in Zusammenarbeit mit dem Technischen Komitee CEN/TC 122 „Ergonomie“ erarbeitet, dessen Sekretariat vom DIN gehalten wird.

Dieses Dokument ist derzeit zur parallelen Umfrage vorgelegt.

Anerkennungsnotiz

Der Text von ISO/DIS 9241-154:2011 wurde vom CEN als prEN ISO 9241-154:2011 ohne irgendeine Abänderung genehmigt.

ISO 9241 besteht aus folgenden Teilen unter dem allgemeinen Titel *Ergonomic requirements for office work with visual display terminals (VDTs)*:

- Teil 1: *General introduction*
- Teil 2: *Guidance on task requirements*
- Teil 3: *Visual display requirements*
- Teil 4: *Keyboard requirements*
- Teil 5: *Workstation layout and postural requirements*
- Teil 6: *Guidance on the work environment*
- Teil 9: *Requirements for non-keyboard input devices*
- Teil 11: *Guidance on usability*
- Teil 12: *Presentation of information*
- Teil 13: *User guidance*
- Teil 14: *Menu dialogues*
- Teil 15: *Command dialogues*
- Teil 16: *Direct- manipulation dialogues*
- Teil 17: *Form filling dialogues*

ISO 9241 besteht ebenfalls aus den folgenden Teilen unter dem allgemeinen Titel *Ergonomics of human-system interaction*:

- Teil 20: *Accessibility guidelines for information/communication technology (ICT) equipment and services*
- Teil 100: *Introduction to standards related to software-ergonomics* [Technical Report]
- Teil 110: *Dialogue principles*

- *Teil 129: Guidance on software individualization*
- *Teil 143: Form-based dialogues*
- *Teil 151: Software ergonomics for Guidance on World Wide Web user interfaces*
- *Teil 154: Design guidance for interactive voice response (IVR) applications*
- *Teil 171: Guidance on software accessibility*
- *Teil 210: Human-centred design for interactive systems*
- *Teil 300: Introduction to electronic visual display requirements*
- *Teil 302: Terminology for electronic visual displays*
- *Teil 303: Requirements for electronic visual displays*
- *Teil 304: User performance test methods for electronic visual displays*
- *Teil 305: Optical laboratory test methods for electronic visual displays*
- *Teil 306: Field assessment methods for electronic visual displays*
- *Teil 307: Analysis and compliance test methods for electronic visual displays*
- *Teil 308: Surface-conduction electron-emitter displays (SED) [Technical Report]*
- *Teil 309: Organic light-emitting diode (OLED) displays [Technical Report]*
- *Teil 310: Pixel defects — Visibility, aesthetics and ergonomics [Technical Report]*
- *Teil 400: Principles and requirements for physical input devices*
- *Teil 410: Design criteria for physical input devices*
- *Teil 420: Selection procedures for physical input devices*
- *Teil 910: Framework for tactile and haptic interaction*
- *Teil 920: Guidance on tactile and haptic interactions*

Zu weiteren Teilen in Vorbereitung siehe Anhang A.

Einleitung

Dieser Teil der ISO 9241 enthält Festlegungen zu Sprachdialogsystemen (IVR, en: interactive voice response), die eine Kombination von Sprachtechnologien umfassen können, sich jedoch durch die Anwendung des Telefons als Informationsübertragungsmechanismus auszeichnen. Diese Festlegungen setzen keine visuellen Informationsanzeigen außer der Beschriftung auf den Tastenfeldern des Telefons voraus, mit der Ausnahme von Texttelefonen (TTY), die ein visuelles Anzeigenfeld für die Rückmeldung haben.

Obwohl es für Gestalter von Benutzungsschnittstellen für Sprachdialogsysteme äußerst wichtig ist, die kulturellen und linguistischen Gesichtspunkte der Benutzungsschnittstelle zu berücksichtigen, die sich auf die vorgesehene Bevölkerungsgruppe auswirken, fallen diese Gesichtspunkte nicht mehr in den Anwendungsbereich dieses Teils der ISO 9241 und werden nicht behandelt. Ebenso wenig fallen detaillierte Festlegungen für die allgemeine Gestaltung von Benutzungsschnittstellen mit automatischer Spracherkennung (ASR, en: automatic speech recognition) in den Anwendungsbereich dieses Dokuments, da automatische Spracherkenner in verschiedenen Sprachen unterschiedlich funktionieren und die Technik sich weiter verbessert. Stattdessen befasst sich dieser Teil der ISO 9241 mit der Gestaltung der Benutzungsschnittstellen von Sprachdialogsystemen und nur mit den Bereichen der Gestaltung von ASR-Benutzungsschnittstellen, die sich auf die Dialoggestaltung auswirken.

Viele gegenwärtige Sprachdialogsysteme stellen signifikante Zugangsschwierigkeiten für behinderte Menschen dar. Einige der Festlegungen in diesem Teil der ISO 9241 wurden insbesondere für Anrufer mit besonderen Bedürfnissen entwickelt, insbesondere gehörlose Menschen oder Menschen mit Hörschäden.

Es ist vorgesehen, dass die Festlegungen in diesem Dokument mit ISO/IEC 13714, *Information technology - Document processing and related communication – User interface to telephone-based services - Voice messaging applications*, kompatibel sind, außer wenn anderweitig angegeben.

Interaktive Sprachdialogsysteme (IVR) sind seit den späten 1980er Jahren ein verbreitetes Mittel im Kundenservice. Diese Systeme wurden gestaltet, um die Notwendigkeit von Telefonschleifen bei der Kundenbetreuung zu reduzieren oder zu beseitigen, indem viele der Funktionen automatisiert wurden, die üblicherweise von Menschen bei der Kundenbetreuung am Telefon bei der Bearbeitung von Kundentransaktionen geleistet wurden. Somit können Benutzer (d. h. Anrufer) Tätigkeiten wie das Abfragen von Zugfahrplänen, Buchbestellungen, Berichte über Probleme mit dem Kabelfernsehen durch Interaktion mit einem Sprachdialogsystem ausführen. Zusätzlich haben Sprachdialogsysteme häufig eine automatische Anrufweiterleitungsfunktion, damit der Anrufer mit der richtigen Stelle für seine spezifische Anfrage verbunden werden kann.

Dieser Teil der ISO 9241 beschäftigt sich weitgehend mit der Gestaltung des Dialogs zwischen Mensch und Sprachdialogsystem. Wie Bild 1 zeigt, interagieren Benutzer mit dem Sprachdialogsystem üblicherweise entweder über Spracheingabe oder über Tonwahl (DTMF) über das Telefontastenfeld. Wenn ein Sprachdialogsystem sprachgesteuert ist, verwendet es eine automatische Spracherkennungskomponente (ASR), die die Spracheingabe des Anrufers erkennt. Wenn es nicht sprachgesteuert ist, erkennt es üblicherweise nur Tonwahleingaben über das Telefontastenfeld oder manchmal Texttelefon-Eingaben (TTY). Sprachgesteuerte Sprachdialogsysteme sind eine relativ neue Entwicklung, und viele Systeme werden nun so gestaltet, dass sie beim Dialog mit einem Anrufer sowohl Tonwahl als auch Sprache erkennen. Das Sprachdialogsystem reagiert über Hardware und/oder Software, die dem Anrufer synthetische, digitalisierte oder aufgezeichnete Sprache vorspielt und zusätzlich nicht-sprachliche akustische Ausgaben ermöglicht. Die Tatsache, dass bei diesen Anwendungen keine visuelle Informationsanzeige vorhanden ist, stellt eine Herausforderung für die Dialoggestalter dar, weil das Navigieren durch die Anwendung sowie die Verarbeitung und das Erinnern der relevanten Informationen ohne Hilfe einer visuellen Anzeige eine Belastung für den Anrufer bedeutet.

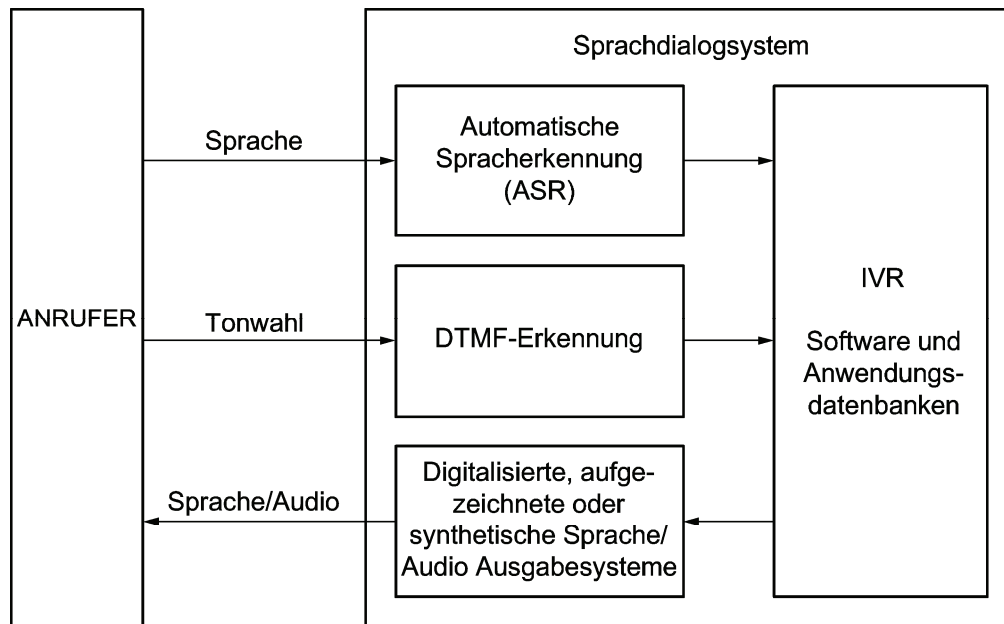


Bild 1 — Darstellung eines Sprachdialogsystems

Dieser Teil der ISO 9241 gibt Leitsätze zur Gestaltung von IVR-Dialogen in sprachgesteuerten und auf Tonwahl beruhenden Sprachdialogsystemen. Deshalb liegt der Schwerpunkt dieses Teils der ISO 9241 auf der Schnittstelle zwischen Benutzer und IVR-Software sowie Datenbanken der Anwendung, die von Hardware und Software vermittelt wird, die Sprache und/oder Tonwahl erkennt und Sprachausgaben ermöglicht. Sowohl ASR-Systeme als auch DTMF-Erkennung bewirken Einschränkungen bei der Gestaltung von Sprachdialogen, und diese Einschränkungen wurden bei der Entwicklung der Leitsätze dieses Teils der ISO 9241 berücksichtigt. Dennoch ist aus zwei Gründen nicht vorgesehen, dass dieser Teil der ISO 9241 sich der Gestaltung von ASR oder DTMF-Benutzungsschnittstellen widmet. Erstens entwickelt sich die Technologie insbesondere für ASR-Systeme immer noch weiter und zweitens sind viele ASR-Benutzungsschnittstellen so gestaltet, dass sie eine visuelle Anzeige enthalten (z. B. ASR-Diktatanwendungen), über die IVR-Anwendungen nicht verfügen. Zuletzt sind die meisten Leitsätze für sprachgesteuerte Dialoge zur Anwendung mit auf Grammatiken beruhenden Sprachdialogsystemen vorgesehen. Obwohl einige derselben Prinzipien für Systeme mit natürlichem Sprachverstehen gelten (NLU, en: natural language understanding, d. h. Anwendungen, die statistische Sprachmodelle verwenden), sind ausführliche Gestaltungsleitsätze für diese Anwendungen in dieser Norm nicht enthalten, da natürliche Spracherkennung mit einer anderen Technologie durchgeführt wird, und die Anwendung natürlicher Sprache in sprachgesteuerten Dialogsystemen sich noch in der Entwicklung befindet. Ebenso gibt es einige Gesichtspunkte der Anwendungsgestaltung, die für natürliche Sprachdialoge hinsichtlich der auf Grammatik beruhenden Sprachdialoge verschieden sind, zusätzlich zu den Unterschieden in der Dialoggestaltung, die sich besonders auf Spracherkennung beziehen.

1 Anwendungsbereich

Dieser Teil der ISO 9241 enthält Festlegungen und Anforderungen zur Gestaltung von Benutzungsschnittstellen von Sprachdialogsystemen. Er behandelt sowohl Sprachdialogsysteme die das Tonwahlverfahren verwenden sowie die Systeme, die automatische Spracherkennung als Eingabeform verwenden. Er gilt gleichermaßen für Fälle, in denen der Benutzer oder das Sprachdialogsystem selbst (z.B. in einigen Fernsehmarketinganwendungen) den Anruf initiieren.

Dieser Teil der ISO 9241 ist zur gemeinsamen Anwendung mit ISO/IEC 13714 vorgesehen.

ANMERKUNG Dieser Anwendungsbereich ist daher allgemeiner als der von ISO/IEC 13714, die spezifisch für Mitteilungsanwendungen.

2 Normative Verweisungen

Die folgenden zitierten Dokumente sind für die Anwendung dieses Dokuments erforderlich. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

ISO/IEC 13714, *Information technology — Document processing and related communication — User interface to telephone-based services — Voice messaging applications*

ITU-T E 161, *Arrangement of digits, letters and symbols on telephones and other devices that can be used for gaining access to a telephone network*

3 Begriffe

Für die Anwendung dieses Dokuments gelten die folgenden Begriffe.

3.1

Ansage

Mitteilung des Sprachdialogsystems an den Anrufer, die informiert, jedoch den Anrufer nicht zu Handlungen anleitet

ANMERKUNG Einige Industriestandards unterscheiden nicht zwischen Ansagen und Prompts und benennen sämtliche vom System ausgegebenen Nachrichten als Prompts.

[ANSI/HFES 200.4]

3.2

automatische Spracherkennung

ASR

(en: automatic speech recognition)

Umwandlung gesprochener Worte in eine computerlesbare Eingabe

ANMERKUNG Dies wird manchmal als automatisierte Spracherkennung bezeichnet.

3.3

Barge-In-Funktion

Fähigkeit eines Sprachdialogsystems, eine Eingabe anzunehmen, während ein Prompt (oder eine Ansage) abgespielt wird

ANMERKUNG 1 Das Abspielen von Sprache endet sofort und das System antwortet auf die Eingabe des Benutzers.

ANMERKUNG 2 Siehe Dial-Through-Funktion oder Talk-Through-Funktion. Der Begriff „Barge-In-Funktion“ ist bei Systemen mit Tastaturwahleingabe gleichbedeutend mit dem Begriff „Dial-Through-Funktion“. Er ist ebenfalls gleichbedeutend mit Talk-Through-Funktion bei Systemen, die Spracheingabe annehmen.

3.4

Anrufer

Benutzer, der einen Dienst anruft, mit dem Sprachdialogsystem verbunden wird und mit diesem über ein Anwendungsprogramm interagiert

ANMERKUNG Der Anrufer initiiert die Interaktion mit eine Dienst oder Sprachdialogsystem.

[ANSI/HFES 200.4]

3.5

Coaching

Anweisungen für die Sprecher der Prompts und Ansagen (auch Sprecherstimme genannt) zu gewünschten Feinheiten bei der Aufzeichnung der Prompts

3.6

verknüpfte Prompts

verknüpfte Eingabeaufforderungen

Prompts oder Ansagen, die durch das Aneinanderreihen mehrerer einzelner Prompts oder Ansagen erstellt wurden

ANMERKUNG Verknüpfte Prompts werden üblicherweise für die Darstellung dynamischer oder kontextabhängiger Information verwendet.

3.7

Konfidenzschwellen

Gewissheit, mit der das Spracherkennungssystem ein Erkennungsergebnis für eine erhaltene Eingabe wiedergibt

ANMERKUNG 1 Siehe Zuverlässigkeitswertung.

ANMERKUNG 2 Konfidenzschwellen werden häufig über bestimmte Bereiche von Konfidenzwerten festgelegt, üblicherweise Hoch, Mittel und Niedrig. Diese Bereiche werden dann verwendet, um bestimmte Sprachdialogantworten auszulösen, z. B. Annahme, Zurückweisung oder Bestätigung der Spracheingabe des Anwenders.

3.8

Konfidenzwert

vom ASR-System zugewiesener numerischer Wert, der das Maß der Gewissheit angibt, mit der die Äußerung des Anrufers richtig erkannt wurde

ANMERKUNG Das ASR-System gibt für jede Nutzereingabe eine Liste möglicher Kandidatenwörter oder Sätze mit einem numerischen Wert zurück, der die angenommene Wahrscheinlichkeit der Korrektheit repräsentiert.

3.9

kontinuierliche Spracherkennung

Spracherkennung, die eine kontinuierliche Eingabe von Wörtern oder Sätzen zulässt

ANMERKUNG Es ist keine Mindestsprechpause zu Beginn oder am Ende von Wörtern oder Sätzen erforderlich (außer Pausen, wie sie im Allgemeinen bei normativen Sprachmustern vorkommen). Der Gegensatz dazu ist die „diskrete Spracherkennung“.

[ANSI/HFES 200.4]

3.10

dialogorientierte Berichtigung

Fähigkeit eines ASR-Systems, eine Gesprächsstörung zu beheben, die aufgrund eines Sprach- oder Tonwahleingabefehlers vorkommt, indem die nachfolgende Spracheingabe des Anwenders verwendet wird, um den geeigneten nächsten Schritt im Dialog festzulegen

BEISPIEL Der Sprecher bei einer Reiseveranstaltungsanwendung sagt „Buffalo“, wenn er nach seinem Reiseziel gefragt wird. Das Sprachdialogsystem antwortet: „Sie möchten nach Chicago, richtig?“ In der Antwort entgegnet der Anrufer: „Nein, Buffalo.“ Wenn das System dialogorientierte Berichtigung verwendet, kann es feststellen, dass der Anrufer eine „berichtigende“ Eingabe gemacht hat, die Eingabe erkennen und dazu verwenden, den Dialog weiterzuführen. Die Antwort des Systems könnte nun sein: „Oh, Sie meinten Buffalo, Verzeihung. Wann möchten Sie abreisen?“, anstatt dem Anrufer dieselbe Frage erneut zu stellen (d. h. „Wohin möchten Sie reisen?“).

3.11

Löschungsfehler

Kategorie eines Erkennungsfehlers, bei der ein Teil der Nutzeräußerung im Ergebnis der Spracherkennung fälschlicherweise fehlt

3.12

Dial-Ahead-Funktion

Fähigkeit eines Sprachdialogsystems, eine Tonwahleingabe anzunehmen, bevor das System dazu aufgefordert hat

ANMERKUNG 1 Die Tonwahleingabe wird dann vom System auf der Grundlage der Reihenfolge verwendet, in der die Eingabe empfangen wurde. Das ermöglicht Anrufern, Eingaben zu machen, ohne die entsprechenden Prompts hören zu müssen.

ANMERKUNG 2 Siehe Talk-Ahead-Funktion.

[ANSI/HFES 200.4]

3.13

Dial-Through-Funktion

Fähigkeit eines Tonwahlbasierten Sprachdialogsystems, die Eingabe eines Anrufers anzunehmen, während ein Prompt (oder eine Ansage) abgespielt wird

ANMERKUNG 1 Siehe Talk-Through-Funktion und Barge-In-Funktion.

ANMERKUNG 2 Als Rückwirkung auf die Dial-Through-Funktion wird die Sprachabspielung angehalten und das System reagiert auf die gedrückte Taste.

3.14

digitalisierte Sprache

digitalisierte Aufzeichnung menschlicher Sprache

ANMERKUNG Steht im Gegensatz zu synthetischer Sprache, aufgezeichneter Sprache und Text/Sprache-Umwandlung.

[ANSI/HFES 200.4]

3.15

Hinweiswörter

Wort, Satz oder Ton, der als Hinweis für den Anrufer verwendet wird, dass ein neuer Prompt oder eine neue Ansage beginnt oder dass vom Anwender nun eine Eingabe erwartet wird

ANMERKUNG Typische Hinweiswörter sind „okay“, „in Ordnung“, „jetzt“.

3.16

diskrete Spracherkennung

Spracherkennung, die eine Sprechpause mit einer Mindestdauer zu Beginn und am Ende eines zu erkennenden Wortes oder Satzes erfordert, damit eine geeignete Analyse der Spracheingabe möglich ist

ANMERKUNG Steht im Gegensatz zu „kontinuierlicher Spracherkennung“.

3.17

Dualtonmultifrequenz

DTMF

Tastaturtöne heutiger Telefontastenfelder

[ANSI/HFES 200.4]

3.18

dynamische Grammatik

für die Spracherkennung verwendete Grammatik, die nicht fest vordefiniert ist

ANMERKUNG Sie wird üblicherweise in Echtzeit auf der Grundlage veränderlicher Daten erstellt.

BEISPIEL Ein Anrufer wird gebeten, eine vorgegebene Nummer zum Ersetzen zu sagen und die Grammatik wird nur auf der Grundlage der vorgegebenen Nummer des Anrufers erstellt, nicht aus sämtlichen möglichen numerischen Kombinationen.

3.19

End-Pointing

Endmarken

Orientierungspunkt

Erfassen von Beginn und Ende einer Spracheingabe

3.20

Enrollment

Training der Spracherkennung

Verfahren, in dem Benutzer Trainingsäußerungen eingeben, um ein sprecherabhängiges Spracherkennungssystem zu trainieren

[ANSI/HFES 200.4]

3.21

explizite Bestätigung

Verfahren, bei dem vom Anrufer verlangt wird, seine letzte Eingabe in ein Sprachdialogsystem zu bestätigen

3.22

Grammatik

syntaktisches und manchmal morphologisches Regelwerk, das sämtliche Äußerungen des Anwenders festlegt, die das Spracherkennungssystem abdeckt

ANMERKUNG Die Grammatik legt die Eingaben fest, die von der Anwendung erkannt werden.

3.23

implizite Bestätigung

Verfahren, die Eingabe eines Anwenders in ein Sprachdialogsystem zu bestätigen, wobei das System auf die Eingabe so reagiert, als ob sie richtig wäre

ANMERKUNG 1 Im Fall der impliziten Bestätigung weiß der Anrufer aufgrund der folgenden Aktion des Sprachdialogsystems, welche Eingabe empfangen wurde.

ANMERKUNG 2 Implizite Bestätigungen sind eine natürliche Art und Weise, den Dialog zu beschleunigen. Jedoch ist der Mechanismus für die Wiederherstellung des Dialogs im Fehlerfall (Widerruf) nicht klar.

BEISPIEL 1 Wenn eine weibliche Anruferin „Kontostand“ angibt, um anzugeben, dass sie vom System den Stand ihres Bankkontos zu erfahren wünscht, zeigt folgende Antwort des IVR „Ihr Kontostand liegt bei \$ 492,19“ implizit an, dass das System die Anfrage nach Informationen zum „Kontostand“ richtig verstanden hat, im Gegensatz zu anderen Angaben zum Konto.

BEISPIEL 2 Anrufer: „Morgen“. – Sprachdialogsystem: „Um welche Zeit möchten Sie morgen abreisen?“

3.24

Einfüigungsfehler

Kategorie eines Erkennungsfehlers, bei der ein oder mehrere Wörter des Ergebnisses der Spracherkennung mit keinem Wort (oder keiner Wortfolge) der Nutzeräußerung übereinstimmen

3.25

interaktiver Sprachdialog

IVR

(en: Interactive Voice Response)

Softwareanwendung, mit der ein Anwender über eine Telefonleitung interagiert, und die vorher aufgezeichnete und/oder dynamisch erzeugte Sprachausgaben umfasst und die Tonwahl- und/oder Spracheingaben des Anrufers entgegen nehmen kann

3.26

Orientierungspunkt

⟨IVR⟩ kurzer Ausdruck wie „Bankkontoinformation“ oder „Reparatordienste“ oder ein Ton, der die Überschrift oder den Anfangspunkt für einen Teil des IVR-Dialogs darstellt und zur Erleichterung der Navigation innerhalb einer IVR-Anwendung für den Anwender verwendet werden kann

3.27

Mitteilung

⟨IVR⟩ Information innerhalb eines Sprachdialogsystems, die vom System, von anderen Anrufern, anderen Teilnehmern oder Systemadministratoren für den Anrufer erstellt wird

ANMERKUNG Mitteilungen umfassen sowohl Prompts (d.h. Handlungsanweisungen) als auch Ansagen (d. h. es ist keine Handlung erforderlich).

[ANSI/HFES 200.4]

3.28

gemischte Initiative

⟨IVR⟩ Dialogfluss, bei dem die Initiative sowohl vom Anrufer als auch vom System ausgehen kann

3.29

natürliches Sprachverstehen

NLU

(en: natural language understanding)

Technologie, die verwendet wird, um bestimmte Wörter und Sätze einer Anruferäußerung zu erkennen, die gesprochen ist, als würde zu einem anderen Menschen gesprochen werden

ANMERKUNG NLU versteht den Anrufer nicht tatsächlich, sondern wird mit anderer Technologie kombiniert, um die Bedeutung der Anruferaussage abzuleiten.

3.30

offener Dialog

Dialog, bei dem die sprachlichen Antworten eines Anwenders nicht eingeschränkt werden

ANMERKUNG Das wird üblicherweise gemeinsam mit NLU angewendet.

BEISPIEL „Was kann ich für Sie tun?“