



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

Computerlinguistik

Lehrinheit 8b: Speech Technology Applications for
the Broad Public

Dr. Christina Alexandris
Nationale Universität Athen
Deutsche Sprache und Literatur

Σκοποί ενότητας

Παρουσίαση Παραδειγμάτων /
Συστημάτων από Έργα



Speech Technology Applications for the Broad Public

The CitizenShield Dialog System
Πιλοτικό Διαλογικό Σύστημα ΠΟΛΙ.ΑΣ

www.polias.gr

Brief Description of the CitizenShield Project

- The Citizenshield Project involves a hybrid approach combining key-word recognition and free input for managing spoken user-input entered in the **CitizenShield** (“POLIAS”) system for consumer complaints.
- The Citizenshield dialogue system developed is targeted towards an improvement of the service of the EKPIZO consumer organization **by relieving the organization’s call center of routine tasks** and, subsequently, allowing the operators to use their time and energy in handling complex cases.



Processing Spoken Input

- The spoken input of the customer's complaints is automatically entered into templates containing a number of fields related to the categories and types of information concerning the product involved.
- The **automatic filling-in of complaint forms** with spoken input via the organization's call center is especially helpful to mobile users and users that have no internet access.
- In the **hybrid approach** concerned, spoken input is either processed by a key-word recognition mechanism or directly entered as free input in the respective field in the complaint form.
- The type of processing used is related to the type of question produced.



User Requirements and Data Processing

- Based on **data** collected from conducted research concerning user requirements, the dialogue system supports **directed dialogs**
- and allows the management of the three most commonly occurring customer complaint categories, namely food products, defective or problematic retail goods and banking and insurance services.
- The series of questions involved constitute a directed dialog, created from a combination of requirements related to speech processing efficiency and the EKPIZO organization requirements and policy.



Processable Complaint Forms

- The information contained in each field of the complaint form is automatically or manually processable, according to the type of task to be executed.
- These complaint forms constitute a **processable data base** from which on-line information is generated, for example, graphs depicting percentages of product types or companies involved in the majority of current customer complaints.



Requirements of the Input Management for the Speech Recognition Module – A Brief Description

Input Management for the Speech Recognition Module

- 1. Using Keyword Lists as Input Management for Speech Recognition
- The multilingual usability features of the Keyword Lists to be developed for the Speech Recognition Component are linked to user input control, in the form of keyword-groups.
- Although keyword-group user-input may vary in respect to the user, this type of input cannot deviate considerably from being restricted and hence, manageable for multilingual applications and allowing minimum interference of language-specific factors.



1. Using Keyword Lists as Input Management for Speech Recognition

- The broad-public usability features of the Keyword Lists to be developed for the Speech Recognition Component are linked to user input control, in the form of keyword-groups.
- Although keyword-group user-input may vary in respect to the user, this type of input cannot deviate considerably from being restricted and hence, manageable for broad-public applications and allowing minimum interference of user-specific factors.



2. Organisation of Keyword Categories

Elements recognized by the system may be divided into three main categories:

- Elements consisting of keywords that are mapped in respect to closed and relatively small lists.
- Elements that are mapped in respect to open databases, such as names, (addresses and locations may be added)
- Elements consisting of numbers that may include information such as quantity, address or date.

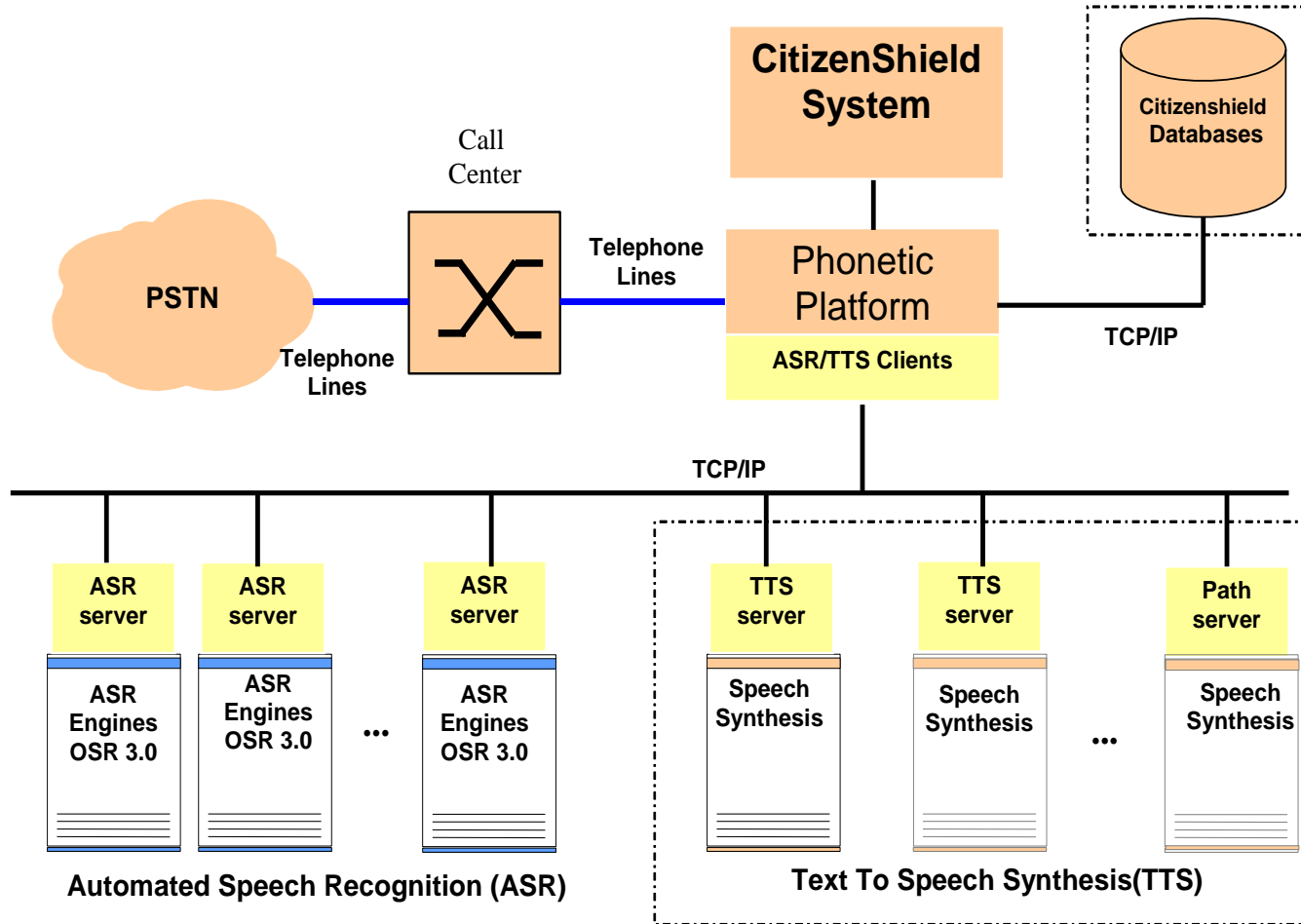


The effort must be made for the utterances produced by the system to be

- (1) clear and unambiguous towards the user
- but at the same time to be (2) friendly and natural-sounding
- and, in addition, to contain expressions that, from a semantic aspect, (3) constrain the range of the user's possible responses to a minimum,
 - thus restricting as far as possible the probability of ambiguities and misinterpretations regarding user input.



CitizenShield Dialog System – Overview (Nottas et al., 2007)



[1]



Prosodic Modeling

The acceptable prosody to the target user-groups, constituting the general public in the System is modelled with the use of emphasis as prosodic marker, placed on the appropriate elements of the utterance for achieving

(A) user-friendliness in man-machine communication in the sense of “accuracy” and “directness” (Hausser 2006) as well as

(B) a sense of control to the user for the CSh System (Alexandris 2007, Nottas et al. 2007) and

(C) familiarity as an additional feature, according to the Project’s User Requirements.



Evaluation

The evaluation of the Project was primarily based on Interaction Aspects used for testing the performance of the System, namely

(1) the Utterance,

(2) Functional

and (3) Satisfaction Levels (Moeller 2005)

and also due both to the complexity of the System

(a) user-group, a remarkable subset of which constitutes senior citizens, according to the Project User Requirements (Work-Package 1) and the

(b) relatively positive results of the evaluation of the systems prosodic modelling in the Evaluation Phase (Work-Package 7)



Evaluation of the implementation of the prosodic modeling of the Conversational Agent's utterances-I

Interaction Aspects related to :

(1) Utterance Level, (2) Functional Level
and (3) Satisfaction Level (3) (Moeller, 2005)

Evaluation performed according to Degree of Naturalness / Acceptability: (“informativeness”, “intelligibility” (Moeller, 2005))



Evaluation of the implementation of the prosodic modeling of the Conversational Agent's utterances-II

Results of Evaluation Phase 50-Users Group:

80%: “Very natural-sounding” (40%)/ “quite natural-sounding” (40%) and 20% “satisfactory results”

- Evaluation performed according to Degree of Credibility: (“perceived task success”, “functional limits” (Moeller, 2005))

Results of Evaluation Phase 50-Users Group:

50%: “Seems to inspire very high credibility” (33,3%) / Seems to inspire quite high credibility” (16,6%) and 50%: «satisfactory results»



Examples of Word Category I from the CitizenShield Dialog System (prosody and interpretation of sentence)

Examples of preferred utterances by Users (prosodic emphasis is underlined) :

[5]: INTERACTION 5: SYSTEM: Please answer the following questions with a «yes» or a «no». Was there a problem with the products packaging?

[USER: YES/NO]

[6]: INTERACTION 6: SYSTEM: At what price did you buy the product?

[USER: PRICE (EURO)]

[10]: INTERACTION 10: SYSTEM: Please tell us any additional information you wish about the product or about your transaction.

[USER: FREE INPUT]

[5.3.1]:INTERACTION 5: SYSTEM: How did you realize this? Please speak freely. Please confirm to me that you are finished. As soon as you are ready, please press «1».

[USER: FREE INPUT]

(Translation from Modern Greek with proximity to original syntactic structure)



Quantifiers and Numericals

– Presence of Prosodic Emphasis = indexical interpretation (=“exactly”)

- Περιμένετε δύο λεπτά (wait for two minutes)
- INFORMATION: TEMPLATE:WAIT
- TIME (2 MINUTES)

– Absence of Prosodic Emphasis = fixed expression

- Περιμένετε δύο λεπτά (wait for a few minutes) (= fixed expression)
- INFORMATION:TEMPLATE:WAIT
- TIME (FEW MINUTES)

Prosodic emphasis (underlined) used for ambiguity resolution (Examples chosen from Alexandris 2007a, Alexandris 2005).



Discourse Particles used as Politeness Markers

Discourse particles (*in italics*) as politeness markers:

– **CITIZENSHIELD CONVERSATIONAL AGENT:**

- *Tell me, what-is the product (for example, milk, coffee, something else)*
- *Πείτε μου τι είναι το προϊόν (παραδείγματος χάρη, γάλα καφές, κάτι άλλο)*



System Evaluation - Examples

- Η Αρχική Αξιολόγηση του Πιλοτικού Διαλογικού Συστήματος ΠΟΛΙ.ΑΣ πραγματοποιήθηκε την συνδρομή συνόλου εθελοντών, ανδρών και γυναικών, προερχομένων από ένα μεγάλο φάσμα ηλικιών και ευρύ πεδίο επαγγελματικής ή άλλου είδους (νοικοκυρές, συνταξιούχοι) απασχόλησης. Επιπλέον, δόθηκε η δυνατότητα στα μέλη της ΕΚΠΟΙ.ΖΩ να παρέχουν τα σχόλιά τους στην ιστοσελίδα www.polias.gr.
- Η Αξιολόγηση πραγματοποιήθηκε σε δύο φάσεις, την Α' Φάση Αξιολόγησης και την Β' Φάση Αξιολόγησης.



3. Β' Φάση Αξιολόγησης

3.1.2 Κριτήρια Αξιολόγησης Β' Φάσης

- Ως προς τα Κριτήρια Αξιολόγησης, στην Β' Φάση Αξιολόγησης δόθηκε έμφαση στην ποιότητα της Αναγνώρισης Φωνής και συγκεκριμένα, την απόδοση του Συστήματος σε σχέση με τις ακόλουθες παραμέτρους:
- **θορυβώδες περιβάλλον** (δρόμος, καφετέρια, ίντερνετ καφέ)
- **περιπτώσεις στις οποίες ο Χρήστης αποκλείει από την προβλεπόμενη μορφή διαλόγου**, τόσο ως προς την καθαρότητα εκφοράς των λέξεων (2.1) όσο και ως προς την χρήση γνωστών και άγνωστων λέξεων από το Σύστημα (2.2).
- Επιπλέον, αξιολογήθηκαν και (3) **τεχνικά θέματα** στην ποιότητα της διαδικτυακής και τηλεφωνικής σύνδεσης του Συστήματος.



3.2 Αποτελέσματα

3.2.1 Ποιότητα Αναγνώρισης Φωνής και απόδοση του Συστήματος σε θορυβώδες περιβάλλον (1/2)

Παρατηρήθηκε ότι οι λέξεις που αναγνωρίζονται από το Σύστημα, τόσο σε θορυβώδες περιβάλλον (δρόμος, καφετέρια, ίντερνετ καφέ), όσο και σε ήσυχο περιβάλλον μπορούν να διαχωριστούν σε δύο κατηγορίες, στις **λέξεις που αναγνωρίζονται σε κάθε περίπτωση (Κατηγορία A)** και σε **λέξεις που αναγνωρίζονται είτε μόνον σε ήσυχο περιβάλλον, είτε καθόλου (Κατηγορία B)**.



3.2.1 Ποιότητα Αναγνώρισης Φωνής και απόδοση του Συστήματος σε θορυβώδες περιβάλλον (2/2)

Στο ακόλουθο παράδειγμα παρουσιάζονται ενδεικτικές κατηγορίες λέξεων της **Κατηγορίας A** που αναγνωρίζονται από το Σύστημα, τόσο σε θορυβώδες περιβάλλον, όσο και σε ήσυχο περιβάλλον.

*Παραδείγματα λέξεων **Κατηγορίας A**:*

- **Προϊόν:** γιαούρτι, φακές, σπαράγγια
- **Εταιρείες:** Τρία Άλφα , Ξενία, La Vache qui rit
- **Ποσότητα:** ένας κεσές, ένα σακκουλάκι, ένα κουτί



Απόδοση Συστήματος σε θορυβώδες περιβάλλον -2

- Η **Κατηγορία Β** περιλαμβάνει κατά κύριο λόγο
 - (i) **λέξεις ξένης ονομασίας** ή «ξενικά» ονόματα, όπως, για παράδειγμα, «γκούντα» (τυρί), «έμμενταλ» (τυρί), «κραφτ» (εταιρεία), «ελ» (elle) (ελληνική εταιρεία), (αν και αναγνωρίζονται ονόματα εταιρειών όπως, για παράδειγμα, «λα βας κι ρί» - La Vache qui rit»)
 - (ii) αρκετές **κατηγορίες μη-τυποποιημένων προϊόντων** όπως, για παράδειγμα, «μπακλαβαδάκια» και
 - (iii) **λιγότερο γνωστές εταιρίες** όπως, για παράδειγμα, «Χανιωτάκη».



3.2.2 Απόδοση Συστήματος: Απόκλιση από προβλεπόμενη μορφή διαλόγου

Ως προς την απόδοση του Συστήματος σε περιπτώσεις απόκλισης Χρήστη από προβλεπόμενη μορφή διαλόγου, παρατηρήθηκε ότι η **απόδοση του Συστήματος δεν είναι καλή** όταν:

- ο Χρήστης **βιάζεται να μιλήσει** με το Σύστημα ή **εκφέρει περισσότερα από ένα στοιχεία**, όπως, για παράδειγμα, «ένα σακουλάκι εε.. μιά συσκευασία» ή
- ο Χρήστης **δεν μιλάει αρκετά δυνατά** /«ζωηρά»/με «σβησμένη» φωνή (κούραση, κακή διάθεση κ.α.).

Οι περιπτώσεις αυτές κατηγοριοποιούνται ως «**Περίπτωση Β**».

- Οι **περιπτώσεις στις οποίες δεν ισχύουν οι περιπτώσεις αυτές** ως μορφές απόκλισης Χρήστη από προβλεπόμενη μορφή διαλόγου κατηγοριοποιούνται ως «**Περίπτωση Α**».



3.2.3 Τεχνικά θέματα στην ποιότητα της διαδικτυακής και τηλεφωνικής σύνδεσης του Συστήματος

- Από την Β΄ Φάση Αξιολόγησης του Συστήματος προκύπτει ότι τυχόν προβλήματα στην απόδοση του Συστήματος ως προς την **διαδικτυακή και τηλεφωνική σύνδεσης του Συστήματος** προκύπτουν κατά κανόνα από **εξωτερικούς τεχνικούς παράγοντες** (π.χ. κακή λήψη).



References (1/2)

- Alexandris, C. 2007. Show and Tell: Using Semantically Processable Prosodic Markers for Spatial Expressions in an HCI System for Consumer Complaints. *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*. 4552/2007, 13-22.
- Alexandris, C. Fotinea, S-E and Efthimiou, E. 2005. Emphasis as an Extra-Linguistic Marker for Resolving Spatial and Temporal Ambiguities in Machine Translation for a Speech-to-Speech System involving Greek. In *Proc. of the 3rd International Conference on Universal Access in Human-Computer Interaction (UAHCI 2005)*. Las Vegas, Nevada, USA.
- Alexandris, C., Fotinea, S-E. 2004. Discourse Particles: Indicators of Positive and Non-Positive Politeness in the Discourse Structure of Dialog Systems for Modern Greek, *International Journal for Language Data Processing Sprache und Datenverarbeitung*, 1-2/2004, 19-29.



References (2/2)

- Hausser, R. 2006. A Computational Model of Natural Language Communication, Interpretation, Inference and Production in Database Semantics. Berlin, Springer.
- Moeller, S. 2005. Quality of Telephone-Based Spoken Dialogue Systems. New York, Springer.
- Nottas, M., Alexandris, C, Tsopanoglou, A. and Bakamidis, S. 2007. A Hybrid Approach to Dialog Input in the CitizenShield Dialog System for Consumer Complaints. In Proc. of Human-Computer Interaction (HCI) 2007. Beijing, China.
- Schilder, F. and Habel, C. 2001. From Temporal Expressions to Temporal Information: Semantic tagging of News Messages. In Proc. of ACL-2001, Workshop on Temporal and Spatial Information Processing, 1309-1316, Pittsburgh, Pennsylvania, USA.



Τέλος Ενότητας

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Σημειώματα

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.0.

Έχουν προηγηθεί οι κάτωθι εκδόσεις:

- Έκδοση διαθέσιμη εδώ. <http://eclass.uoa.gr/courses/GS158/>



Σημείωμα Αναφοράς

Copyright Εθνικών και Καποδιστριακών Πανεπιστημίων Αθηνών, Χριστίνα Αλεξανδρή. «Υπολογιστική Γλωσσολογία. Speech Technology Applications for the Broad Public». Έκδοση: 1.0. Αθήνα 2014. Διαθέσιμο από τη δικτυακή διεύθυνση: <http://opencourses.uoa.gr>



Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.



Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.



Σημείωμα Χρήσης Έργων Τρίτων (1/2)

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Εικόνες/Σχήματα/Διαγράμματα/Φωτογραφίες

Εικόνα 1: CitizenShield Dialog System-Overview (Nottas, M., Alexandris, C, Tsopanoglou, A. and Bakamidis, S. 2007. A Hybrid Approach to Dialog Input in the CitizenShield Dialog System for Consumer Complaints. In Proc. of Human-Computer Interaction (HCI) 2007. Beijing, China)



Σημείωμα Χρήσης Έργων Τρίτων (2/2)

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Πίνακες

