

Detecting Linguistic HCI Markers in an Online Aphasia Support Group

Yoram M Kalman
The Open University of Israel
yoramka@openu.ac.il

Kathleen Geraghty, Cynthia K Thompson,
Darren Gergle
Northwestern University
k-geraghty, ckthom, dgergle
@northwestern.edu

ABSTRACT

Aphasia is an acquired language disorder resulting from trauma or injury to language areas of the brain. Despite extensive research on the impact of aphasia on traditional forms of communication, little is known about the impact of aphasia on computer-mediated communication (CMC). In this study we asked whether the well-documented language deficits associated with aphasia can be detected in online writing of people with aphasia. We analyzed 150 messages (14,754 words) posted to an online aphasia support forum, by six people with aphasia and by four controls. Significant linguistic differences between people with aphasia and controls were detected, suggesting five putative linguistic HCI markers for aphasia. These findings suggest that interdisciplinary research on communication disorders and CMC has both applied and theoretical implications.

Categories and Subject Descriptors

H.1.2 [Information Systems]: Models and Principles – *human factors*

General Terms

Measurement, Human Factors.

Keywords

Aphasia, computer-mediated communication, HCI markers, human factors, online support groups, unobtrusive monitoring, user modeling.

1. INTRODUCTION

HCI markers are signals created during human-computer interaction (HCI) which might provide information about the cognitive, mental, psychological or physiological state of the user [24]. The concept of HCI markers is analogous to the concept of biomarkers, which are indicators of physiological or disease processes. Some widely known biomarkers used for health assessment include physical measures such as body temperature or blood pressure, as well as biochemical measures such as hormone levels. Changes in these biomarkers have been linked to a diverse set of conditions including cancer, Alzheimer's disease, metabolic disorders, stress, and cardiovascular disease [12, 14, 17, 21]. In

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ASSETS'12, October 22–24, 2012, Boulder, Colorado, USA.
Copyright 2012 ACM 978-1-4503-1321-6/12/10...\$15.00

this paper, we study messages posted by users with aphasia who contribute to an online support forum, and present evidence for the presence of HCI markers for aphasia, a language disorder resulting from neurological disease. As suggested in Kalman [24], such HCI-related variables can be identified and associated with important health-related conditions. Furthermore, HCI markers can be used for purposes such as unobtrusive detection, diagnosis, and monitoring of various conditions [e.g. 25, 31], as well as for dynamically personalizing the user interface to the needs of users with impairments [15].

1.1 HCI Markers

Our online activity reveals much about who we are, our interests, and our relationships [3]. It is suggested that our online activities and the way we interact with user interfaces also reveals important information about our health, and that the signals created during human-computer interaction (HCI) can be used to detect health-related changes [24]. In the same way biomarkers such as cholesterol and blood sugar levels are used to detect, diagnose and monitor physiological processes such as cardiovascular disease and diabetes, Kalman suggests identifying HCI markers to detect diagnose and monitor HCI-related variables that are affected by conditions such as neurological damage.

One of the chief candidates for HCI marker analysis is online language. An abundance of research links variables such as personality, emotion and neurological health to language use [e.g. 16], but we are still early in the process of understanding the impact of such conditions on online language [e.g. 9, 13]. The goal of the current study is to further our understanding of the link between online language and neurological damage. We attempt to determine whether language patterns associated with brain damage can be detected in online writing. This is a required first step before making use of this information in beneficial ways in existing and new technologies.

1.2 Aphasia

Aphasia is an acquired language disorder, resulting from brain trauma or injury to the language areas of the brain, which may affect all modes of receptive and expressive communication, including speech, language, writing, comprehension, and gesture [18]. Although there are many shared characteristics of aphasic language, people with aphasia present with varying language deficit profiles within and across language domains. A person's language capacities may remain stable in some types of aphasia, or progressively decline, as is the case in Primary Progressive Aphasia (PPA) [29, 30]. Aphasia is often noted for being a frustrating and often debilitating communicative disorder. Because of the increased probability of stroke and neurological disease in older populations, diagnosis and treatment of aphasia

and other degenerative language conditions is important for increasingly aging populations, such as the Baby Boomer generation.

While there are different types of aphasia, people with aphasia generally present with problems in both production and comprehension of language. In language production, errors may include the omission of key linguistic features used by neurologically intact speakers, and/or disordered use of these features. In addition, they often present with mild to severe problems with language comprehension [2]. In measurable terms, production problems often lead to overall decreases in sentence length (MLU – mean length utterance), omission of word classes (e.g. prepositions), and a range of semantic, syntactic and phonological errors [33-35].

Users with language disabilities experience unique challenges when they interact via user interfaces. In the case of people with aphasia, the challenges include reading from the screen, using voice commands, and writing, and extensive efforts are focused on developing assistive technologies to overcome these challenges [e.g. 1, 11, 36], and even to use computers to increase empathy towards people with aphasia by developing software that emulates the language distortion they experience [19, 20]. However, since little is known about the manifestation of aphasia during computer-mediated communication, the individual attributes of the online communication of people with aphasia are not being used to diagnose and to monitor the condition.

2. RESEARCH QUESTIONS AND HYPOTHESES

This study asks whether the well-documented language deficits associated with aphasia can be detected in online writing. While research on aphasia *per se* is extensive, there is surprisingly sparse research analysis of online interactions of people with aphasia. Although there are many reasons why we might predict patterns similar to spoken language, there are other reasons why online language may reveal distinctive characteristics [7]. Based on the known characteristics of aphasia in traditional (mainly spoken) communication, we chose seven linguistic variables that are typically measured in spoken aphasic language [See 33 for a detailed analysis of linguistic deficits found in individuals with primary progressive aphasia]. We hypothesized that the online writing of people with aphasia would differ significantly when compared to the online writing of people without aphasia (“controls”). We predict that online writing of people with aphasia will exhibit similar language deficits. If this prediction is confirmed, then these patterns are HCI marker candidates.

Specific hypotheses were formulated for each measurable language variable. For variables measuring linguistic complexity, we hypothesized that people with aphasia would rank lower than controls. For variables measuring linguistic errors, we hypothesized that errors would be significantly higher for people with aphasia than controls. The hypotheses and variables were thus:

Linguistic complexity, as measured by (1), (2), and (3) below will be decreased for people with aphasia as compared to controls.

(1) MLU: the mean number of words produced per utterance.

(2) Open/closed class ratio: the ratio of all open class words (i.e., nouns, verbs, adjectives, and adverbs) to all closed class words produced.

(3) Noun/verb ratio: the ratio of all nouns to all verbs produced.

Linguistic errors, (4), (5), (6), and (7) below, will be produced at higher rates for people with aphasia as compared to controls.

(4) Proportion of ungrammatical utterances: the number of utterances with grammatical errors, divided by total number of utterances.

(5) Proportion of morpheme inflection errors: the number of errors in morpheme inflection, divided by total number of inflected morphemes.

(6) Proportion of open class errors: the number of errors in open class words, divided by the number of all open class words.

(7) Proportion of closed class errors: the number of errors in closed class words, divided by the number of all closed class words.

3. METHOD

3.1 Participants

A public online support group for people with aphasia was the source for the messages analyzed in this study. We first read the online posts of the participants in the support group. We selected six participants with aphasia, and four control participants who did not have aphasia, based on self-descriptions from their posts. For analysis, we selected only participants who posted at least 15 messages, and clearly identified their medical status: (1) a person with aphasia, (2) or a person who does not have aphasia. It was not possible for us to ascertain the official medical diagnosis or standardized participant demographics, since their online posts included only self-reported information. However, based on self-report, the age of the youngest person with aphasia was mid-30's and the oldest about 70. Five of the six people with aphasia mentioned that the cause of their aphasia was a stroke. Most reported living in North America, and English as their native language. Some mentioned command of additional languages such as French and Chinese. For controls, three of the four were relatives of a person with aphasia: a spouse, a child and a parent, and one was the spouse of a speech-language pathologist. No further personally identifying details or direct quotes from the support forum are provided, to protect the anonymity of the participants in the support group.

3.2 Language Coding

We developed a protocol for hand-coding online written language based on existing aphasia research protocols, with necessary adaptations [see 33-35]. The protocol included three phases. First, each posted message was segmented into utterances, based on punctuation, other visual demarcations (e.g. new line) and verb use. Second, the part of speech (POS) of each word was determined, based on the Penn Treebank POS parsing framework [28]. Last, each utterance was coded for errors: written convention errors such as spelling and capitalization errors, and grammatical errors: syntactic and morphological. The results of the coding were used for the analysis described below.

3.3 Analysis

The results of the coding were used to calculate: MLU, total word count, utterance count, count of grammatical and ungrammatical sentences, count of nouns and verbs, count of inflected morphemes and count of morpheme inflexion errors, count of each POS (part of speech), count of open class words and errors in open class words, count of closed class words and count of errors in closed class words. These were used to calculate the variables of interest. Because of the small number of participants, we used a conservative nonparametric Wilcoxon signed-rank test to analyze differences between the two participant groups.

4. RESULTS

Table 1 summarizes the results of coding for each of the ten participants. As expected and as is typical with analyses of spoken language in aphasia research, we see high variability within both the people with aphasia, and control groups (e.g. the number of words and utterances in each post). In Table 2, each participant's score for each variable is presented in detail. Again, we see variability between participants, although the controls clearly show lower rates of linguistic errors. Table 3 presents the statistical analysis: results of the Wilcoxon signed-rank test comparing people with aphasia to controls. Our hypotheses were confirmed for five of the seven variables that showed a significant difference ($p \leq .01$) between people with aphasia and controls. MLU was significantly shorter for people with aphasia, and the rates of the four linguistic error types were higher for people with aphasia. There was no significant difference between the open/closed class and the noun/verb ratios.

Table 1. Results by participant

Participant ^a	101	102	103	104	105	106	901	902	903	904
Mean length of utterance	7.68	8.44	9.25	10.77	10.07	6.08	13.77	13.50	16.49	15.26
Words	1928	996	668	1228	3047	260	1526	1070	2011	2020
Utterances	242	123	72	113	317	33	110	82	126	127
Correct utterances	158	87	44	82	214	25	101	77	116	124
Incorrect utterances	84	36	28	31	103	8	9	5	10	3
Inflected morphemes	350	22	93	220	79	50	175	137	185	275
Morpheme inflection errors	20	2	4	15	13	1	3	1	5	1
POS (parts of speech) tags	2057	1065	646	1297	3199	279	1494	1118	2108	2049
Open class words	1355	582	363	748	1775	145	871	689	1143	1209
Open class errors	65	23	20	12	86	5	4	3	5	6
Closed class words	702	483	283	549	1424	134	623	429	965	840
Closed class errors	27	22	12	10	43	9	4	4	8	1
Nouns	404	189	149	196	551	47	259	228	375	465
Verbs	459	193	121	272	666	62	293	246	414	363

^a People with aphasia: Participants 101-106. Controls: Participants 901-904.

Table 2. Language Variables for All Participants

Participant ^a	101	102	103	104	105	106	901	902	903	904
Mean length of utterance	7.68	8.44	9.25	10.77	10.07	6.08	13.77	13.50	16.49	15.26
Ungrammatical sentences	34.71%	29.27%	38.89%	27.43%	32.49%	24.24%	8.18%	6.10%	7.94%	2.36%
Morpheme inflection errors	5.71%	9.09%	4.30%	6.82%	16.46%	2.00%	1.71%	.73%	2.70%	.36%
Open class errors	4.80%	3.95%	5.51%	1.60%	4.85%	3.45%	.46%	.44%	.44%	.50%
Closed class errors	3.85%	4.55%	4.24%	1.82%	3.02%	6.72%	.64%	.93%	.83%	.12%
Ratio, open/closed class	.88	.98	1.23	.72	.83	.76	.88	.93	.91	1.28
Ratio, noun/verb	1.93	1.21	1.28	1.36	1.25	1.08	1.40	1.61	1.18	1.44

^a People with aphasia: Participants 101-106. Controls: Participants 901-904.

Table 3. Comparison of Language Variables: People with Aphasia vs. Controls, and Significance Test Using Wilcoxon Signed-Rank Test

	People with aphasia		Controls		Hypothesis (H_1)	p-value
	Average	SD	Average	SD		
Mean length of utterance	8.72	1.70	14.75	1.39	People with aphasia < Controls	.005
Ungrammatical sentences	31.17%	5.28%	6.14%	2.69%	People with aphasia > Controls	.005
Morpheme inflection errors	7.40%	5.04%	1.38%	1.05%	People with aphasia > Controls	.01
Open class errors	4.03%	1.39%	.46%	.03%	People with aphasia > Controls	.005
Closed class errors	4.03%	1.64%	.63%	.36%	People with aphasia > Controls	.005
Ratio, open/closed class	.90	.19	1.00	.19	People with aphasia > Controls	n.s.
Ratio, noun/verb	1.35	.30	1.41	.17	People with aphasia > Controls	n.s.

5. DISCUSSION

As predicted, people with aphasia showed reduced linguistic complexity and increased presence of errors in their online writing. This demonstrates that some of the major language deficits of aphasia can be detected in online writing of people with aphasia. We discuss the implications of these findings on the study of HCI markers, the implications for practice and for theory, and the limitations of the study.

5.1 Implications for Practice

To the best of our knowledge, this is the first time HCI markers for aphasia are being proposed in the literature. The five variables that differed significantly between people with aphasia and controls unequivocally distinguish between these two groups, and are thus candidate aphasia HCI markers. Extensive additional work needs to be done in order to establish these markers, to evaluate their specificity and sensitivity, and to determine guidelines for their use. A better understanding of these markers has practical as well as theoretical implications, as discussed below.

At present, clinical assessment of aphasia is based on face-to-face interviews, followed by manual interview transcription, and analysis of the transcript by an expert clinician. The establishment of HCI markers for aphasia could augment this practice in several ways. First, additional linguistic samples could now be collected unobtrusively from online language in e-mail correspondence, online posts to social networking sites such as Facebook, blogs, or online groups (health oriented support groups, as well as other special interest groups). The resulting text is already transcribed in digital format making it immediately available for analysis, and it is produced under conditions that are often more naturalistic than the traditional face-to-face clinical evaluation. It is also possible to capture instant messaging sessions, though it is expected that the nature of synchronous communication HCI markers will differ from those described here.

Because clinicians would now be able to collect more extensive and representative language samples from patients with less effort, this could both augment their diagnostic and therapeutic toolset, as well as their ability to monitor longitudinal changes in aphasic language symptoms over periods of months and years. HCI markers could augment current test batteries, as well

as allowing clinicians to explore the progression of language decline prior to arrival at the clinic, especially in the case of ongoing deterioration (e.g. in the case of PPA). Thus, an analysis of online language created and archived pre-diagnosis (e.g. old e-mails, blogs, or old Facebook posts) could help evaluate the rate of linguistic decline as well as its trajectory. As online writing becomes more prevalent in the general population as well as with older adults [26, 27], access to such longitudinal repositories of language samples will improve. In regards to monitoring the progression of aphasia, the same principles apply, and the clinicians would be able to better evaluate either recovery (e.g. from stroke-induced aphasia), or decline (e.g. from PPA), and how these are influenced by various interventions.

One of the more exciting possibilities suggested by our findings is that linguistic HCI markers will be useful in earlier detection of progressive aphasia as well as of other progressive neurodegenerative diseases which influence language use, such as Alzheimer's disease [16]. The realization of this vision requires the development of automated or semi-automated parsing and analysis tools [23] that would be able to monitor, on an ongoing basis, language produced by the users. Despite the challenges of automated language analysis, it should be possible to adapt parsing tools to perform automated or semi-automated identification and quantification of HCI markers.

Another intriguing possibility is applying these findings to linguistic HCI markers to improve assistive technologies offered to people with language disorders [e.g. 1, 11, 36]. Lastly, there is the question of whether participation in online discussion forums is also a worthwhile remedial effort and supplemental to face-to-face therapy sessions. These technologies could ostensibly monitor the putative positive impact of ongoing participation in online conversations and interactions.

5.2 Implications for Theory

In a recent paper by Gajos et al. [15], the authors point to the paucity of data on actual behavior of users with disabilities in real world situations. Data on clients' use of technology in clinical settings may be gathered as a part of diagnostic evaluation, at the discretion of the clinician. These data are necessary in order to realize the concept of personalized dynamic accessibility. HCI markers are ideal candidates for data that will allow measuring and modeling users' abilities. The

same unobtrusively collected HCI markers that can assist in the detection, diagnosis and longitudinal monitoring of specific conditions, are also critical measures of users' abilities to interact with various digital devices. HCI markers for aphasia are especially interesting, since, as Gajos et al. emphasize, modeling cognitive abilities is one of the toughest challenges. The markers identified in this paper are a small step in the direction of developing such modeling capabilities.

One of the main sources for information about human behavior and health is a result of the study of disorders and illness. Just as metabolic disorders provide clues on the structure of metabolic pathways and anxiety disorders shed light on the processing of fear and anxiety, aphasia and other language disorders teach us about language structure and production. A better understanding of the expression of aphasia in naturally occurring online communication can also strengthen efforts to effectively emulate aphasia in online communication [19, 20], and a similar approach can be applied towards the emulation of other communication disorders. Very little attention has so far been paid to the convergence of CMC and communication disorders, with the exception of work on assistive technologies [e.g. 22, 32]. We suggest that these early findings on the impact of aphasia on CMC could lead to a better understanding of CMC, as well as of aphasia.

Spoken and written communication are quite distinct from each other. One of the earliest questions raised in regards to computer-mediated communication (CMC) is how it relates to these two traditional forms of communication [6, 7, 10, 37]. Is e-mail simply a conversation in written form, or is it an expedited letter? An analysis of the impact of language disorders on CMC provides a new lens through which to explore this question. Which of the spoken manifestations of aphasia are preserved in the online medium, and which ones diminish or disappear altogether? Here we can see that despite the effectively unlimited time to compose an asynchronous CMC message, MLU decreased in asynchronous CMC produced by people with aphasia. On the other hand, the noun/verb ratio and open class closed class ratios were not affected.

The findings reported here inform both our understanding of language production, and our understanding of CMC as utilized by populations with language disorders. The use of CMC by these populations has not been studied in the past, and the preliminary findings reported here suggest new questions: Are these populations empowered by CMC [4, 5]? How does CMC influence the way these populations are perceived by others? And, how can these findings inform the design of CMC artifacts and of assistive technologies [8]?

5.3 Limitations

In this paper we report, for the first time, the results of research that links the study of communication disorders with the study of CMC. As can be expected from early stage research, our findings are limited in many ways, which should be addressed by future research. First, only a fraction of people with aphasia have the technical and language skills to use CMC; Also, as already mentioned, the specificity and sensitivity of the proposed HCI markers still need to be established. In addition, the results reported here are based on a small cohort of participants; the medical information we have about them is based on self-report, and is fragmentary and unverified. Thus, for example, it is not possible to determine whether HCI

markers can be used to distinguish between different types of aphasia. Even this fragmentary information points to the extensive heterogeneity of the participant pool. Moreover, it is still impossible to determine the value of markers such as noun/verb ratio and open/closed class ratio as HCI markers. Future research efforts should be devoted to studying a larger cohort of participants and/or studying subjects whose aphasia has been studied and classified, as well as to compare their language to carefully matched controls.

6. ACKNOWLEDGMENT

Our thanks to Kristina Rodriguez for her assistance with language coding. Funding was provided, in part, by National Science Foundation grant #0953943.

7. REFERENCES

- [1] Allen, M., McGrenere, J. and Purves, B. The Field Evaluation of a Mobile Digital Image Communication Application Designed for People with Aphasia. *ACM Trans. Access. Comput.* Article 5 (May 2008). DOI=<http://dx.doi.org/10.1145/1361203.1361208>
- [2] American Speech-Language-Hearing Association *Aphasia*, 2011.
- [3] Amichai-Hamburger, Y. *The social net: understanding human behavior in cyberspace*. Oxford University Press, USA, 2005.
- [4] Amichai-Hamburger, Y., McKenna, K. Y. A. and Tal, S. A. E-empowerment: Empowerment by the Internet. *Computers in Human Behavior*, 24, 5 (2008), 1776-1789. DOI=<http://dx.doi.org/10.1016%2Fj.chb.2008.02.002>
- [5] Barak, A. and Sadovsky, Y. Internet use and personal empowerment of hearing-impaired adolescents. *Computers in Human Behavior*, 24, 5 (2008), 1802-1815. DOI=<http://dx.doi.org/10.1016/j.chb.2008.02.007>
- [6] Baron, N. S. *Always on: Language in an online and mobile world*. Oxford Univ Press, 2008.
- [7] Baron, N. S. Letters by phone or speech by other means: The linguistics of email. *Language and communication*, 18 (1998), 133-170. DOI=<http://dx.doi.org/10.1016%2F0271-5309%2898%2900005-6>
- [8] Bowker, N. I. Understanding barriers to online experience for people with physical and sensory disabilities using discursive social psychology. *Universal Access in the Information Society*, 9, 2 (2010), 121-136. DOI=<http://dx.doi.org/10.1007%2Fs10209-009-0162-3>
- [9] Cohn, M. A., Mehl, M. R. and Pennebaker, J. W. Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science*, 15, 10 (2004), 687. DOI=<http://dx.doi.org/10.1111%2Fj.0956-7976.2004.00741.x>
- [10] Collot, M. and Belmore, N. *Electronic language*. John Benjamins, City, 1996.
- [11] Daemen, E., Dadlani, P., Du, J., Li, Y., Erik-Paker, P., Martens, J.-B., de Ruyter, B. Designing a free style, indirect, and interactive storytelling application for people with aphasia. In *Proceedings of the 11th IFIP TC 13 international conference on Human-computer interaction (INTERACT'07)*. Springer-Verlag, Berlin, Heidelberg, 221-234.
- [12] Diamandis, E. P. Cancer Biomarkers: Can we turn recent failures into success? *Journal of the National Cancer*

- Institute*, 102, 19, 1462-1467. DOI=
<http://dx.doi.org/10.1093%2Fjnci%2Fdjq306>
- [13] Epp, C., Lippold, M. and Mandryk, R. L. Identifying emotional states using keystroke dynamics. In *Proceedings of the CHI 2011* (Vancouver, BC, Canada, (2011), New York. DOI=
<http://dx.doi.org/10.1145%2F1978942.1979046>
- [14] Foundation for the National Institutes of Health *The Biomarkers Consortium*. Foundation for the National Institutes of Health, 2011.
- [15] Gajos, K. Z., Hurst, A. and Findlater, L. Personalized dynamic accessibility. *interactions*, 19, 2 (2012), 69-73. DOI= <http://dx.doi.org/10.1145%2F2090150.2090167>
- [16] Garrard, P. Cognitive archaeology: Uses, methods, and results. *Journal of Neurolinguistics*, 22, 3 (2009), 250-265. DOI=
<http://dx.doi.org/10.1016%2Fj.jneuroling.2008.07.006>
- [17] Gerszten, R. E. and Wang, T. J. The search for new cardiovascular biomarkers. *Nature*, 451, 7181 (2008), 949-952.
- [18] Goodglass, H. *Understanding aphasia*. Academic Press, San Diego, CA, 1993.
- [19] Hailpern, J., Danilevsky, M., Harris, A., Karahalios, K., Dell, G. and Hengst, J. ACES: Promoting empathy towards aphasia through language distortion emulation software. In *Proceedings of the 2011 annual conference on Human factors in computing systems (CHI '11)*. ACM, New York, NY, 609-618. DOI=
<http://doi.acm.org/10.1145/1978942.1979029>
- [20] Hailpern, J., Danilevsky, M. and Karahalios, K. *ACES: aphasia emulation, realism, and the Turing test*. In *The proceedings of the 13th international ACM SIGACCESS conference on computers and accessibility (ASSETS '11)*. ACM, New York, NY, 83-90. DOI=
<http://dx.doi.org/10.1145/2049536.2049553>
- [21] Hellhammer, D. H., Wüst, S. and Kudielka, B. M. Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology*, 34, 2 (2009), 163-171. DOI=
<http://dx.doi.org/10.1038%2Fnature06802>
- [22] Jeon, M., Gupta, S., Davison, B. K. and Walker, B. N. Auditory menus are not just spoken visual menus: a case study of unavailable menu items. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems (CHI EA '10)*. ACM, New York, NY, 3319-3324. DOI=
<http://doi.acm.org/10.1145/1753846.1753978>
- [23] Jurafsky, D. and Martin, J. H. *Speech and language processing: An introduction to natural language processing, computational linguistics and speech recognition*. Prentice Hall, 2008.
- [24] Kalman, Y. M. *HCI markers: A conceptual framework for using human-computer interaction data to detect disease processes*. The 6th Mediterranean Conference on Information Systems (MCIS), Limassol, Cyprus, 2011.
- [25] Madan, A., Cebrian, M., Moturu, S., Farrahi, K. and Pentland, S. Sensing the 'health state' of a community. *Pervasive Computing*, (2011). DOI=
<http://dx.doi.org/10.1109%2FMPRV.2011.79>
- [26] Madden, M. *Older Adults and Social Media*. Pew Research Center, 2010.
- [27] Madden, M. and Zickuhr, K. *65% of online adults use social networking sites*. Pew Research Center, 2011.
- [28] Marcus, M. P., Marcinkiewicz, M. A. and Santorini, B. Building a large annotated corpus of English: The Penn Treebank. *Computational linguistics*, 19, 2 (1993), 313-330.
- [29] Mesulam, M. M. Primary progressive aphasia: a language-based dementia. *New England Journal of Medicine*, 349, 16 (2003), 1535-1542. DOI=
<http://dx.doi.org/10.1056%2FNEJMra022435>
- [30] Mesulam, M. M., Wieneke, C., Rogalski, E., Cobia, D., Thompson, C. K. and Weintraub, S. Quantitative template for subtyping primary progressive aphasia. *Archives of neurology*, 66, 12 (2009), 1545. DOI=
<http://dx.doi.org/10.1001%2Farchneurol.2009.288>
- [31] Moreno, M. A., Jelenchick, L. A., Egan, K. G., Cox, E., Young, H., Gannon, K. E. and Becker, T. Feeling bad on Facebook: depression disclosures by college students on a social networking site. *Depression and anxiety*, 28, 6 (2011), 447-455. DOI=
<http://dx.doi.org/10.1002%2Fda.20805>
- [32] Smyth, K. and Kwon, S. Computer-mediated communication and its use in support groups for family caregivers. *Gerontechnology: Research and practice in technology and aging* (2004), 97-116.
- [33] Thompson, C. K., Ballard, K. J., Tait, M. E., Weintraub, S. and Mesulam, M. M. Patterns of language decline in non-fluent primary progressive aphasia. *Aphasiology*, 11, 4/5 (1997), 297-321. DOI=
<http://dx.doi.org/10.1080%2F02687039708248473>
- [34] Thompson, C. K., Cho, S., Wieneke, C., Weintraub, S. and Mesulam, M. M. Dissociation between fluency and agrammatism in Primary Progressive Aphasia. *Aphasiology* (January, 2012). DOI=
<http://dx.doi.org/10.1080%2F02687038.2011.584691>
- [35] Thompson, C. K., Shapiro, L. P., Tait, M. E., Jacobs, B., Schneider, S. and Ballard, K. A system for the linguistic analysis of agrammatic language production. *Brain and Language*, 51, 1 (1995), 124-129.
- [36] Woudstra, M., Al Mahmud, A. and Martens, J.-B. A snapshot diary to support conversational storytelling for persons with aphasia. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, 641-646. DOI=
<http://doi.acm.org/10.1145/2037373.2037474>
- [37] Yates, S. J. Oral and written linguistic aspects of computer conferencing. *Computer-mediated communication: Linguistic, social and cross-cultural perspectives* (1996), 29-46.