Ingressive speech as an indication that humans are talking to humans (and not to machines)

Robert Eklund
Robert.H.Eklund@telia.se
Telia Research AB, SE–123 86 Farsta, Sweden & Linköping University, IDA/NLP-LAB, Sweden

ABSTRACT

- Speech produced on pulmonic ingressive airstreams (different from velaric and glottalic ingressive speech sounds) is often anecdotally mentioned in the linguistic literature. Most previous studies have stressed the paralinguistic function of ingressive speech (IS). This paper studies IS in two corpora of spontaneous Swedish speech. Eight subjects made business travel bookings in two data collections. In one corpus the subjects talked with what they believed was a computer, played by a professional actor, in the other corpus the same subjects talked with a human travel agent. All subjects made use of ingressive speech in the Human–Human setting, and no subject made use of ingressive speech in the Human–"Machine" setting. The results are discussed from the perspective of possible underlying factors, including discourse structure, gender

1 INTRODUCTION

- Pulmonic ingressive speech (IS) – speech produced on inhaled airstreams – is often anecdotally mentioned in the literature.
- Not many studies have been devoted to IS.
- Often said to have paralinguistic function in human communication.
- Reeves & Nass (1996) claim that humans interact with media in basically the same way they interact with other human beings.
- Implications of this claim are discussed, based on observations in this study,
- Possible enhancements in Human–Machine interaction are suggested.

2 PREVIOUS STUDIES

- Often claimed that IS is typical of Scandinavian languages.
- IS has been reported in:
  - Swedish (Pitschmann 1987; Allwood 1988, 1988, 1982; Landqvist, 2001)
  - Norwegian (Peters 1981; Kobayashi 2001; Shorrocks [forthcoming])
  - Danish (Stølen 1994)
  - Icelandic (Pitschmann 1987)
  - Finnish (Hakulinen 1993)
  - German (Pitschmann 1987)
  - Maine English (Peters 1981)
  - Newfoundland English (Shorrocks [forthcoming])
  - British English dialects (Shorrocks [forthcoming])
- Also other languages, e.g. Kuot of New Ireland, Tohono O'odham, Swiss German, Damin, French, Dutch and others.
- Also reported in whistle languages.
- For a detailed account of the literature, see: Robert Eklund. *Ingressive Speech: What, How, Where, Who, and Why* (in preparation).

3 DATA AND METHOD

- Two corpora of spontaneous Swedish telephone conversations were used.
- In the first corpus, **WOZ2**, forty-six subjects (32M/14F) booked business travels from what they believed was a computer, played by an professional actor. A total of 140 dialogs/3444 utterances were collected. WOZ2 will be referred to as HM (Human–Machine).
- In the second corpus, **Nymans**, eight subjects (6M/2F) – all of whom participated in WOZ2 – performed the same exact tasks as in WOZ2, this time speaking with authentic, human, travels agents. A total of 24 dialogs/1730 utterances were collected. Nymans will be referred to as HH (Human–Human).
- The tasks were given in mainly iconic form in order to avoid linguistic biasing from the tasks sheets. (See Fig. 1.)
- The subjects/agents were sitting in a room with a headset (used for DAT recordings) and a normal landline telephone (used to tap the telephone line). (See Fig. 2.)

Fig. 1: Example of task sheet.

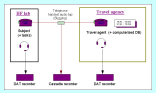

Fig. 2: Nymans setup.

4 RESULTS

4.1 Frequency distribution of ingressive speech items

- All instances of IS are shown in the Table 2 below. (Tables are numbered according to their appearance in the article.)

Table 2: Ingressive speech for clients in WOZ2 (HM) and for clients and agents in Nymans (HH). The question mark indicates that the item is barely audible, even with headphones.

Client	Gender		HM	HH	
	Client	Agent		Client	Agent
1 (LL)	M	F	0	1	0
2 (DM)	M	F	0	2	0
3 (MN)	M	F	0	4	0
4 (FS)	F	F	0	34	12
5 (HS)	M	M	0	12	3
6 (BU)	F	M	0	1	0
7 (RO)	M	M	0	6	0
8 (MS)	M	M	17	8	0
		Σ	17	68	15

- As is shown, while all subjects produced IS in Nymans, no one (with one possible exception, marked with a question mark) produced IS in WOZ2.
- The sole possible example in WOZ2 is barely audible, and probably "inner-directed".
- As a control, the entire WOZ2 corpus (i.e. the 38 subjects who did not participate in Nymans) was checked for IS, without finding a single instance of IS.

4.2 Typology of ingressive speech items

- All instances of IS are typical feedback words, i.e. ingressive variants of (normally egressively produced) words such as "ja" (yes) or "nej" (no), both occurring in voiced and unvoiced form, and affirmative "mm", which is always voiced.
- The distribution is shown in the Table 3 below.

Table 3: Relative frequency of ingressive and corresponding egressive feedback markers pooled for clients and agents in WOZ2 and Nymans. Note that both clients and agents also make use of a variety of other, egressive-only, feedback words.

Type	Client		Agent	
	Ingressive	Egressive	Ingressive	Egressive
ja ('yes')	51	658	14	27
mm ('mm-hm')	12	361	–	–
a ('yeah')	1	3	–	–
nej ('no')	1	52	1	3
jorå ('oh sure')	1	1	–	–
jaså ('indeed')	1	3	–	–
bra ('good')	1	69	–	–
Σ	68	1162	15	30

- The data in Table 3 confirm the literature in that "ja" is by far the most common feedback marker.
- NB! Approximately 10% of all instances of "ja" are produced on ingressive airstream.

4.1 Human–Machine vs. Human–Human.

- Scripted "agent" utterances in WOZ2 included none of the feedback words in Table 3.
- Raises the question whether agent feedback signaling is needed to elicit feedback signaling in the subjects.
- The occurrence of client (subject) produced feedback words, broken down for WOZ2 and Nymans is shown in Table 4.

Table 4: Frequency of feedback words in WOZ2 (HM) and Nymans (HH). As "ja" (yes) counts all variants, such as "jaha".

Client	"ja", "jo" etc.		"mm"		"nej"	
	HM	HH	HM	HH	HM	HH
1 (LL)	4	55	10	28	5	8
2 (DM)	23	69	0	81	9	3
3 (MN)	17	16	9	33	9	3
4 (FS)	11	109	16	121	4	14
5 (HS)	1	112	0	44	2	5
6 (BU)	2	96	0	24	2	7
7 (RO)	1	116	0	34	4	6
8 (MS)	15	143	0	0	4	7
Σ	74	716	35	377	39	52

- Significant differences between the HM and the HH settings.
- Important point: All subjects make use of feedback words in *both* settings.
- Ergo: Agent feedback obviously *not* a prerequisite for feedback in clients/subjects.
- Also: In the four dialogs in Nymans where both agent and client make use of IS, it is the client who initiates its use.

5 DISCUSSION

- Most important observation is that *all* subjects make use of IS in the HH setting, and that *no* subjects make use of IS in the HM setting.
- Potential factors discussed below.

5.1 The role of system feedback

- The lack of system feedback surely to a large degree responsible for the absence of IS in the HM setting.
- Cannot be the full explanation, since the subjects still produced feedback signals in HM.
- Also noted in the literature that IS serves a "bonding" function, not called-for in HM.

5.2 The role of dialog structure

- The agent in WOZ2 used a small set of scripted utterances, while the dialogs in Nymans were fully free.
- Allowed for more genuine speech, which could elicit IS.

5.3 The role of gender

- Often claimed that IS is more employed in female speech.
- Data too scarce to corroborate or refute any assumptions about gender differences.
- IS most common in same-gender dialogs, both noted in the literature and in this study.

5.4 The role of linguistic function

- Most studies mention feedback signaling, but other, more specific, functions are also listed in the literature such as *affective*, *self-directed*, *responsive* and *topic-closing* (Hakulinen); *affirmative*, *inner-directed* and *aligning* (Stølen); *argumentative* and *topic-closing* (Landqvist); *feedback*, *self-reflecting*, *irritation/impatience*, *friendliness/intimacy*, *confirmative*, *concluding* and *turn-taking* (Kobayashi).
- The present data confirm some of the findings in the literature above. The four most common functions are shown in Table 6.

Table 6: Functions of IS in the present corpora. The data for subjects 1–3 and 6–8 are pooled, while the figures for dialog 4 and 5 are broken down for clients and agents. Note that the some figures exceed the sum totals in Table 2, since instances of IS may serve more than one function.

Function	1–3/6–8	4		5	
	Clients	Client	Agent	Client	Agent
Feedback	16	21	7	7	–
Closing	5	7	3	4	–
Intimacy	–	5	5	1	–
Confirming	1	4	3	–	3

- Feedback function by far the most common.
- Closing function second most common function.
- Must be borne in mind that analysis of linguistic function is difficult, and necessarily post-hoc.

6 CONCLUSIONS

- This study confirms Reeves & Nass in that subjects use feedback words in HM setting.
- However, subjects *do* seem to make a difference between machines and humans, using IS only in the HH setting, although the constrained dialogs in WOZ2 are surely partly responsible for these results.
- Could IS be used as a metric to judge the naturalness of a computerized system?
- IS is likely only to appear in natural conversations, but in fully natural automatic systems there could be a need for inclusion in speech recognizer training material.
- Automatic *systems* could also use IS, to serve linguistic, communicative functions.
- Goal (?): To make humans want to "bond with machines".

7 REFERENCES

[1] Allwood, Jens. 1988. The Structure of Dialogue. In: M. M. Taylor, F. Néel & D. G. Bouwhuis (eds.): *The Structure of Multimodal Dialog II*. Amsterdam, John Benjamins, 3–24.

[2] Allwood, Jens. 1988. Om det svenska systemet för språklig återkoppling. In: Per Linell, Mueka Adelswärd, Torbjörn Nilsson & Per A. Pettersson (eds.): *Svenskans beskrivning 15*. Vol. 1. Linköping University. 89–106.

[3] Allwood, Jens. 1982. Finns det svenska kommunikationsnormer? In: *Vad är Lingvistik?* Vol. 7. Papers from the Nordiska Forskarkurser 9. Department of Linguistics, Gothenburg University, 9–40.

[4] Hakulinen, Auli. 1993. Interaktionen som kulturellt interaktionsmönster. In: A.-M. Ivars et al. (eds.): *Språk och social kontroll*. Meddelanden från institutionen för nordiska språk och litteratur vid Helsingfors Universitet, serie B:15. Helsinki University. 49–67.

[5] Kobayashi, Naruki. 2001. *Ingressivt "Ja" på import – ännu ingen på svenskarbete eller dålig hjerne*. MA thesis. Institutt for Lingvistikk og Litteraturvitenskap, Bergen University.

[6] Landqvist, Håkan. 2001. *Från och ruelse. Mönster och samtalsstrategier i tillfrisknandeprocessen inom anonympsykoterapi*. PhD thesis. Department of Scandinavian Languages (Institutionen för Nordiska Språk), Uppsala University.

[7] Peters, Francis Joseph. 1981. *The Paralinguistic Sympathetic Ingressive Affirmative in English and the Scandinavian Languages*. Unpublished PhD thesis, New York University.

[8] Pitschmann, Louis A. 1987. The Linguistic Use of the Ingressive Air Stream in German and the Scandinavian Languages. *General Linguistics*. Vol. 27, No. 3. 153–181.

[9] Reeves, Byron & Clifford Nass. 1996. *The Media Equation. How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge, Cambridge University Press.

[10] Shorrocks, Graham. Forthcoming. Pulmonic Ingressive Speech in Newfoundland English: A Case of Irish-English Influence? In: Hildegard L. C. Tristram (ed.): *The Celtic Englishes III*. Anglistische Forschungen, Heidelberg, Universitätsverlag C. Winter.

[11] Stølen, Marianne. 1994. Gender-related use of the ingressive /a/ in informal conversation among native Copenhagen adults. In: Elizabeth Dahl, Mette-Marie Dyhr & Cathrine Kirby (eds.): *Collected Papers from the Proceedings of the Fifth Berkeley Women and Language Conference*, Berkeley Women and Language Group, Berkeley, 658-677.