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

Robert Eklund
Telia Research JAB, 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 sounds. Here, we discuss possible underlying factors in the distribution of ingressive speech in the Human–Machine setting. The results are discussed from the perspectives of possible underlying factors, including discourse structure, gender, and system feedback.

1 INTRODUCTION
1.1 Pulmonic ingressive speech (IS) — speech produced on inhaled airstreams — is often anecdotally mentioned in the literature. Not many studies have been devoted to IS.
1.2 Often said to have paralinguistic function in human communication. Parsons & New (1995) claim that humans interact with media in basically the same way they interact with other human beings.
1.3 The lack of system feedback surely to a large degree responsible for the absence of IS in the HH setting.

2 PREVIOUS STUDIES
2.1 Often claimed that IS is typical of Scandinavian languages.
2.2 IS has been reported:
   - Norwegian (Hakulinen 2001; 2002; 2003)
   - Dutch (Bekker 1978)
   - Italian (Peters 1981)
   - French (Peters 1981; 1982)
   - Danish (Peters 1981)
   - German (Peters 1981; 1982)
   - Finnish (Hakulinen 1993)
   - Norwegian/English (Hakulinen 1993)
   - British English dialects (Hakulinen 1993)
   - Also, other languages, e.g., Dutch (Hakulinen 1993), Russian (Dekker 2007), Swiss German (Darian 2007; Dutch [343]), French (Darian 2007; Dutch [344]).
   - Also reported in unrelated languages.

3 DATA AND METHOD
3.1 Two corpora of spontaneous Swedish telephone conversations were used.
   - The first corpus, WOZ1, consisted of 3444 utterances from six agents in six dialogues/3444 utterances were collected. WOZ1 will be referred to as HH (Human–Human).
   - The second corpus, Nymans, eight subjects (HH1–HH8) of whom participated in WOZ2 — performed the same exact task as in WOZ1. The data had been collected from authentic, human, travel agents. A total of 1730 utterances were collected. Nymans will be referred to as HM (Human–Machine).

4 RESULTS
4.1 Frequency distribution of ingressive speech forms
   - All instances of IS found in the Table 1 below. (Tables are numerical according to their appearance in the article.)
   - As shown, while all subject produced IS in Nymans, no one, with one possible exception, made use of a question mark produced IS in WOZ2.
   - The wide possible example in WOZ2 is barely audible and possible ‘word/phrase’.
   - As a result, the entire WOZ2 dataset (i.e., the 38 subjects who did not participate in Nymans) was checked for IS, without finding a single instance of IS.

4.3 Typology of ingressive speech forms
   - All instances of IS are bias toward feedback words, i.e., ingressive syllables of (1) ingressive glottalic words (notably speech marker), and (2) suffixing in Swedish, for example, ‘om’ (i.e., the occurence of client produced feedback words, broken down for WOZ2 and Nymans is shown in Table 4.)
   - The distribution is shown in the Table 3 below.

4.4 Human-Machine vs. Human-Human
   - Scipiet’s agent utterances in WOZ2 included none of the feedback words in Table 3. However, the human corpus, WOZ1, only feedback words were classified as ‘feedback’ in Tables 1 and 2 below. The data in Table 3 confirm the literature in the ‘yes’ is for the most common feedback marker.
   - ANOVA: Approximately 95% of all instances of ‘IS’ are produced on ingressive airstream.

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

6 CONCLUSIONS
6.1 This study confirms the results in the literature that feedback words in HH settings.
   - However, subjects do seem to make a difference between machines and humans, using IS only in the HH setting, although the constrained dialogue in WOZ2 are surely partly responsible for these results.
   - Could be used as a metric to judge the robustness of a system?
   - IS is likely only to appear in natural conversations, but in fully natural automatic systems.
   - Could be used as a metric to judge the robustness of a system?
   - How is IS likely to appear in natural conversations, but in fully natural automatic systems.
   - Could be used as a metric to judge the robustness of a system?
   - The occurrence of client produced feedback words, broken down for WOZ2 and Nymans is shown in Table 4.

6.2 Role of gender
   - Reeves & Nass (1996) claim that humans interact with media in basically the same way they interact with other human beings.
   - Data too scarce to corroborate or refute any assumptions about gender differences.
   - IS most common in same-gender dialogues, both noted in the literature and in this study.

6.4 Role of linguistic function
   - Most studies report feedback phrases, but other, more specific, functions are also attested in the literature such as affirmative, negative, smoothing, and controlling.
   - The occurrence of client produced feedback words, broken down for WOZ2 and Nymans is shown in Table 4.

7 REFERENCES