Acoustic analysis of the Estonian encounters data & comparison with DigiSami conversation data.

# 3. DigiSami Corpus and its Annotations

The DigiSami Corpus of spoken North Sami has been collected in the areas traditionally inhabited by the Sami people: in Enontekiö, Utsjoki, Inari and Ivalo in Finland, and in Kautokeino and Karasjok in Norway (see the map in Figure 1). North Sami belongs to the Fenno-Ugric language family, and is one of the Sami languages spoken in Northern Scandinavia, Finland and Kola Peninsula (Seurujärvi et al., 1997). The corpus includes read and conversational speech, and the conversations are both recorded and videotaped. All the speakers are native speakers of North Sami, and their age vary between 16 and 65 years. The data are thus versatile, including informants from two different countries and of different ages. See more about the data collection in Jokinen (2014), Jokinen and Wilcock (2014).

Figure 1. The Sami languages and the data collection [Map]

Although the DigiSami conversational data is not huge (195 minutes of annotated data; mean duration of the conversations was 10:07 minutes), it is valuable because it is the first North Sami conversational corpus. Conversations concern the participants’ everyday life, and their styles differ depending on the age of the speaker and their social status. The topics among young students concern the next vacation, driving school, and cars, while two adult men, mutually acquainted with each other, converse about Sami translation and other technological tools for writing North Sami. The conversations between a pupil and a teacher are fairly formal, and the topics stick to the forthcoming task, i.e. things that one could write a Wikipedia article about.

For the purposes of measuring engagement and to see how laughs function as part of conversations, we annotated the data with laughter features using Praat. Following the previous research, the laughter annotation included the markers for the two laughter types free laugh (fl) and speech-laugh (st), and for the more specific characterization: ‘m’ – mirth, ‘e’ – embarrassed, ‘b’ – breath, ‘p’ – polite, ‘d’ – derision and ‘o’ – other. Table 1 presents the laughter types with explanations.

|  |  |  |
| --- | --- | --- |
| fl | free laughter | laughter without speaking simultaneously |
| st | speech-laugh | laughter and speech combined |
| b | breath | heavy breathing, smirk, sniff; unvoiced, glottal sounds and sibilants |
| e | embarrassed | speaker is embarrassed, confused, uncertain; disassociating |
| m | mirth | fun, humorous, real laughter, occurring when telling jokes etc. |
| d | derision | mocking the partner |
| p | polite | polite laughter showing positive attitude towards the other speaker |
| o | other | laughter that doesn't fit in the previous categories; acoustically unusual laughter |

Table 1. The annotated laughter types (move to Laughter types section?)

The total number of laughter occurrences was 341 in 8 different conversations. Two of these conversations were recorded in Karasjok, Norway, and the rest in Ivalo and Utsjoki, Finland. There were 19 conversation informants altogether – some of the conversations had 2 and some 3 participants. 11 of the participants were female and 8 were male. Altogether, 59% (201) of the laughter occurrences were performed by a female informant, while 41% (140) were performed by a male informant. Table 2 shows the number of different laughter types in different conversations.

# 4. The Estonian encounters corpus and its Annotations

The Estonian encounters data was collected in ???. All participants were adult and in many conversations, the topic was mainly studies or work. None of the participants knew each other beforehand, as opposed to the DigiSami data, in which all participants knew each other or were even close friends. All conversations were the same, rather formal style: the participants started with introducing themselves, then briefly discussed about the conversation situation, how they should stand in front of the camera and what to talk about, then moved on to talking about each others’ studies and career.

The laughter occurrences in the Estonian encounters corpus were annotated accordingly as described above. The total number of laughter occurrences in the Estonian encounters data was 519 in 23 different conversations. All conversations had 2 participants and the mean duration of the conversations was 5:52. 22 of the participants were female and 24 male – each participant attended 2 different conversations with a different partner (one participant had 2 different partners). Altogether, 67% (348) of the laughter occurrences were performed by a female informant, while 33% (171) were performed by a male informant.

# 5. Laughter types

[DigiSami]

The basic statistics are shown in Table X. Free laughter occurs 58% of the laugh occurrences while speech laugh occurs 42% (see discussion below). Three of the specific laughter types occur significantly more frequently than the other types: mirth 29%, embarrassed 49%, and breath 19%, of the total occurrences, and can be called basic laughter types. The laughter bouts annotated as derision, polite and other together only account for 3% of the total occurrences, and can be considered marked types of laughter.

The differences between different conversations can be seen when the laugh activity is normalized with respect to the time. In our data, the average number of laughs is 4.8 per minute, but this varies from almost three times more in 02\_V to almost one eight in 07\_SX. Qualitative analysis of the conversations shows that the frequency and types of laughter are linked to how well the participants know each other, how nervous they are, and what kind of relationship they have with each other. For example, in the conversation 02\_V in which the participants laugh and chuckle the most, they know each other very well, whereas in 07\_SX where only a handful of laughs occur, the speakers’ relationship is

asymmetrical and the whole interaction more formal.

When studying the most laugh-active and engaging conversation 02\_V more closely, we notice that the relative count of free laughter is 79% and that of speech-laugh 21%, i.e. the percentage of free laughter is almost four times more than speech-laughs. A closer analysis shows that half of the laughter instances are mirthful or embarrassed laughs, and the other half breathy sounds. This is in contrast with the other conversations where laughter seems to be either mirthful or breathy.

It appears that 02\_V is an exception among the conversations in other respects, too: its free laughs account for two thirds of the free laughs in the whole data and it also has most of the embarrassed laughter occurrences. In fact, laughing in 02\_V seems to function quite unlike laughing in the other conversations: it signals uncertainty, confusion or embarrassment. This is supported by observations that conversation topics change very fast and have long silences in them, and that the speakers seem nervous in general.

Ignoring laughs in 02\_V, we notice that free laughing is reduced to only one third of all the laughing occurrences (60/167), i.e. laughter simultaneously with speech seems to be more common than free laughing, and can obviously be used as an effective signal of the speaker’s engagement and attitude. On the other hand, 02\_V seems to exemplify that laughing is also used as an effective strategy to relieve stress and confusion, besides indicating the speaker’s personal characteristics and conversational roles.

In general, we can hypothesise that in natural conversations where people know each other and show no overt nervousness, the basic laughter types occur in two situations: when the participants have real fun, i.e. when telling jokes or funny stories, or when they provide breathy feedback to the partner to signal their engagement in the conversation. However, if the conversational situation creates nervousness, this can be signalled by two extremes: by excessive laughter, or by lack of laughter. The former is common among peers who can thus jokingly share their confusion, uncertainty and embarrassment, while the latter is common among strangers and participants who have asymmetrical power relations and thus markedly signal their non-sharing: laughing automatically creates closeness and in-group feeling which makes the partners more equal.

The basic statistics of the Estonian Encounters data are shown in Table X. Free laughter occurs 57% of the laugh occurrences while speech laugh occurs 43%. Two of the specific laughter types occur significantly more frequently than the other types: mirth 36% and breath 54%, of the total occurrences, and can be called basic laughter types. The laughter bouts annotated as embarrassed, polite and other together only account for 9% of the total occurrences, and can be considered marked types of laughter. There was no laughter bouts annotated as derision in the Estonian Encounters data. Compared to the DigiSami data, the amount of embarrassed laughter was significantly lower in the Estonian Encounters data. Instead, there were several occurrences of polite laughter in situations where the participants wanted to seem encouraging to one another.

In the table below are shown the durations of each conversation and the amounts of different laughter types.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Conv. | Duration | fl | st | m | b | e | p | o | Tot. | /min |
| C\_1 | 5:50 | 18 | 13 | 10 | 14 | 7 | 0 | 0 | 31 |  |
| C\_2 | 6:19 | 9 | 13 | 9 | 10 | 0 | 3 | 0 | 22 |  |
| C\_3 | 6:08 | 12 | 4 | 7 | 8 | 0 | 1 | 0 | 16 |  |
| C\_4 | 6:20 | 17 | 3 | 0 | 12 | 3 | 5 | 0 | 20 |  |
| C\_5 | 6:33 | 26 | 15 | 16 | 15 | 7 | 3 | 0 | 41 |  |
| C\_6 | 6:18 | 18 | 22 | 27 | 11 | 2 | 0 | 0 | 40 |  |
| C\_7 | 6:46 | 11 | 21 | 8 | 24 | 0 | 0 | 0 | 32 |  |
| C\_8 | 6:49 | 9 | 9 | 8 | 6 | 3 | 1 | 0 | 18 |  |
| C\_9 | 6:51 | 15 | 7 | 8 | 13 | 0 | 1 | 0 | 22 |  |
| C\_10 | 5:04 | 4 | 13 | 4 | 13 | 0 | 0 | 0 | 17 |  |
| C\_11 | 5:42 | 19 | 18 | 23 | 13 | 0 | 1 | 0 | 37 |  |
| C\_12 | 5:44 | 11 | 11 | 12 | 9 | 1 | 0 | 0 | 22 |  |
| C\_13 | 6:51 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 4 |  |
| C\_14 | 6:07 | 11 | 7 | 3 | 15 | 0 | 0 | 0 | 18 |  |
| C\_15 | 5:52 | 5 | 2 | 1 | 6 | 0 | 0 | 0 | 7 |  |
| C\_16 | 5:28 | 4 | 3 | 1 | 6 | 0 | 0 | 0 | 7 |  |
| C\_17 | 5:18 | 13 | 11 | 12 | 12 | 0 | 0 | 0 | 24 |  |
| C\_18 | 4:43 | 23 | 15 | 15 | 22 | 0 | 1 | 0 | 38 |  |
| C\_19 | 5:24 | 14 | 14 | 5 | 20 | 0 | 3 | 0 | 28 |  |
| C\_20 | 5:01 | 11 | 6 | 2 | 11 | 0 | 0 | 4 | 17 |  |
| C\_21 | 5:12 | 7 | 3 | 8 | 1 | 0 | 1 | 0 | 10 |  |
| C\_22 | 5:38 | 17 | 7 | 4 | 19 | 0 | 1 | 0 | 24 |  |
| C\_23 | 5:08 | 19 | 5 | 4 | 20 | 0 | 0 | 0 | 24 |  |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Conv. | Duration | fl | st | m | b | e | d | p | o | Tot. | /min |
| 02\_V | 13:33 | 138 | 36 | 28 | 87 | 59 | 0 | 0 | 0 | 174 | 13.14 |
| 06\_PS | 11:44 | 0 | 10 | 4 | 4 | 0 | 0 | 2 | 0 | 10 | 8.70 |
| 05\_TP | 8:33 | 18 | 37 | 27 | 27 | 0 | 0 | 0 | 1 | 55 | 6.60 |
| 08\_VV | 12:29 | 12 | 22 | 19 | 15 | 0 | 0 | 0 | 0 | 34 | 2.76 |
| 01\_S | 16:29 | 25 | 18 | 14 | 20 | 6 | 2 | 0 | 1 | 43 | 2.64 |
| 04\_S | 3:28 | 3 | 4 | 2 | 3 | 1 | 0 | 0 | 1 | 7 | 2.16 |
| 03\_V | 8:02 | 2 | 12 | 6 | 8 | 0 | 0 | 0 | 0 | 14 | 1.74 |
| 07\_SX | 6:44 | 0 | 4 | 0 | 2 | 0 | 0 | 2 | 0 | 4 | 0.60 |
| Total | Mean dur. | 198 | 143 | 100 | 166 | 66 | 2 | 4 | 3 | **341** |  |
| % | 10:07 | 58 | 42 | 29 | 49 | 19 | 3 | | | **100** |  |

Table X. The Estonian Encounters data.

# 6. Acoustic analysis

Following the previous research, we hypothesize that the different laughter types in our data differ in their acoustic properties, such as pitch, formants and intensity, and also duration. In the following acoustic analysis, only the basic and most common laughter types, mirthful (m), breath (b) and uncertain/embarrassed (e) and polite (p) in the EE data are included, occurring either in free laughter (fl) or in speech-laughs (st). The analyses have been made with different Praat scripts and further processed for min/max/ave/std values.

[Kuvaaja # of occurrences]

EstEnc # DigiSami #

Our analysis of the acoustic features showed various differences in the investigated laughter types. To compare the results of the acoustic analyses, we calculated average values from the 3-4 most common types (breath, embarrassed, mirth, polite) of male and female informants separately.

The most common laughter types are shown in Figures X and X. As can be seen, in both corpora female participants produce more laughter signals than men, in the DS data usually about twice as many. An exception is free laughing breath types where the ratio is the other way round: this is the typical laugh type for the men in the DS data. It is also interesting that females produce embarrassed and uncertain speech-laughs about four time as many as male participants, being the most typical laugh-type for women in the DS data.

In the Estonian Encounters (EE) data, the most common laughter type is the breathy type both in free laugh and speech-laugh, as shown in Figure X. Also in the EE data, female participants produced remarkably more laughter than male participants.

[Compare /minute ratios?]

Figure X and X shows f0 (pitch) values which were extracted with different ranges for male (75-400Hz) and female (100- 500Hz) informants; thus comparison of male and female average values is not adequate. However, in the DS data it was clear that the f0 in free laughter types was higher than f0 in speech-laughter for both male and female informants, which accords e.g. with Truong and van Leeuwen (2007). In the EE data the polite free laugh had the highest pitch of all laughter types, both for female and male participants. The free laugh types of male participants were clearly higher than the speech-laughs. Generally, there were no big differences between f0 values of all laughter types.

EstEnc F0 DigiSami F0

[Average F0 values]

Figure 4 depicts the average duration of the laugh types. In the DS data, there were big differences in duration between the laugher types: durations of embarrassed laughs were significantly longer (2.1s – 3.2s) than all other types for both male and female informants, and breath laughs were the shortest (1.1s –1.4s). The durations in the EE data were very different compared to the DS data: the longest laughter bout in the EE data was only 1.3 seconds, while in the DS data the longest was 3.2 seconds. One explanation for this significant difference could be the different interaction situation: the EE participants met each other for the first time, while in the DS data the participants knew each other beforehand, thus the laughter bouts were longer.

The EE data had the longest laughter bouts in mirthful types, while the polite and breathy types were the shortest. Both in EE and DS data, the embarrassed speech-laugh was the longest produced by both female and male participants.

However, our data did not support the findings of Nwokah et al. (1999) since free laughter in our data was not significantly longer than speech-laughs. This may be due to the different interaction activities: our data records people conversing in fairly equal situations compared with a mother and child care-giving interaction.

[Average duration]

EstEnc Durations DigiSami Durations

Both in the DigiSami and Estonian Encounters data the **intensity** of different laughter types was rather similar in all laughter types, as shown in Figures X and X. No significant differences occurred between the different laughter types, but the most surprising difference in the DS data was that the average intensity with female informants was generally bigger than with males, while in the EE data many laughter types had higher intensity with male informants.

EstEnc int DigiSami int

# 7. Conclusion

In this paper we have studied the interlocutors’ laughing in the North Sami and Estonian interactions and the functions of laughter in conversational engagement. We can conclude that laughter has several functions that range from fun and happiness to a relief burst of embarrassment. We observed that laughter types depend on the situation and the role of the speakers, and we hypothesize that in natural conversations, the basic laughter types occur when the participants have real fun (mirth) or when they give breathy feedback (breath). However, if the situation is embarrassing or uncomfortable to the speaker, this is signaled by two extremes of laughter frequency, which differ depending on the participants’ power relation: the peers use excessive laughter so as to share the embarrassing situation with the others, whereas the partners in an asymmetrical relation indicate more formal, non-sharing behavior by the lack of laughter.

Also our hypothesis concerning the acoustic differences of different laughter types was supported. In the DigiSami data, durations of embarrassed laughs were significantly longer than the durations of all other types for both male and female informants, but we did not get support for Nwokah et al. (1999) finding of speech-laughs being longer than free laugh. As for the distinctions in intensity, this was small between the laugh types, but interestingly the women had higher intensity laughs than the men in our DigiSami data.

The Estonian first encounters had different laughter types compared to the DigiSami data: it had less embarrassed laughter, but a remarkable amount of polite laughter. In general, breathy and mirthful laughter were the majority of all laughter bouts. We hypothesize that the differences in the laughter types occur mainly because of very different relationship between the participants (first encounters in EE vs. mutually acquainted in DS) and of the age of the participants (all adult in EE vs. partly high school students in DS).

The comparison of acoustic analyses of our datasets showed the most remarkable differences in durational features of the different laughter types: The EE data had the longest laughter bouts in mirthful types, while the polite and breathy types were the shortest. Both in EE and DS data, the embarrassed speech-laugh was the longest produced by both female and male participants. Interestingly, the longest laughter bout in the EE data was only 1.3 seconds, while in the DS data the longest was 3.2 seconds, which is over twice as long as in the EE data.

The other two acoustic features, the f0 and intensity, showed no remarkable differences. In the DS data it was clear that the f0 in free laughter types was higher than f0 in speech-laughter for both male and female informants, which accords e.g. with Truong and van Leeuwen (2007), while in the EE data the polite free laugh had the highest pitch of all laughter types, both for female and male participants.

In both datasets the the **intensity** of different laughter types was rather similar, but the most surprising difference in the DS data was that the average intensity with female informants was generally bigger than with males, while in the EE data many laughter types had higher intensity with male informants.

These observations will be substantiated with deeper statistical analysis, and models for joking, laughing and generally positive attitude will be explored further so as to enable appropriate models be implemented in the SamiTalk application (Wilcock et al. 2016). A useful case is e.g. to be able to recognize the user’s embarrassment or uncertainty on the basis of the amount of laughter and their role in the conversation, and alleviate such situations appropriately.

Moreover, as the collected data is multimodal, it is possible to study non-verbal as well as verbal communication. As argued in the previous research, the participants’ engagement in the conversation and mutual bonding can be measured with the help of multimodal and non-verbal cues, such as the number of laughs or chuckles, or overlapping speech (Bonin, 2016). Our future studies concern the use of non-verbal information when laughing, to measure the participants’ engagement in the interaction.

**8. Acknowledgements**

The work has been conducted within the Academy of Finland project Finno-Ugric Digital Citizens (grant n°270082). The **Praat scripts were made by Mietta Lennes, Shigeto Kawahara (intensity), and Jonas Lindh (F0) and we are grateful for their open code.**