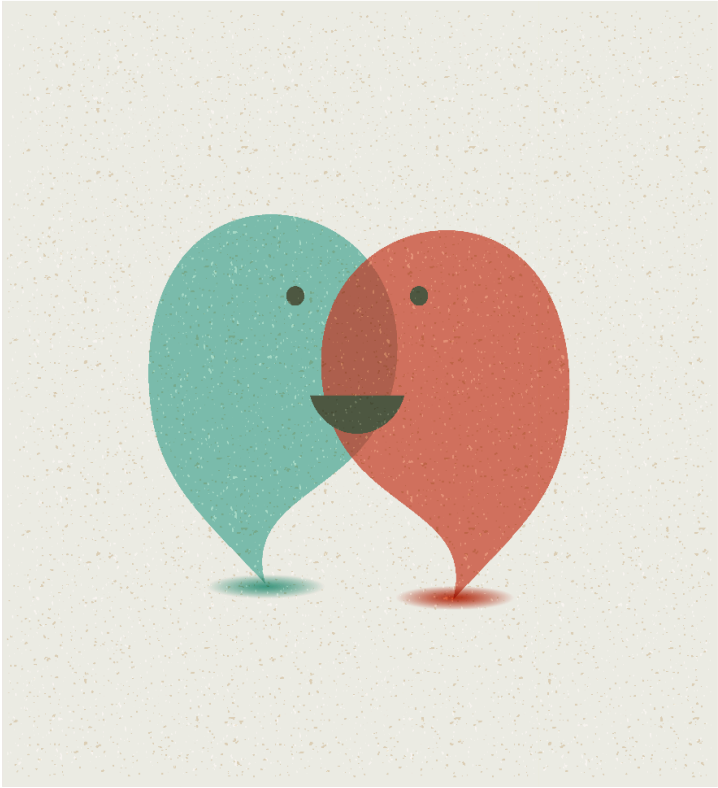
INN.NO

Investigating fluency variables in learner language

Methodological concerns



Hege Larsson Aas & Susan Nacey

Learner Corpus Research Conference
October 2017

INLAND NORWAY UNIVERSITY OF APPLIED SCIENCES

Background: Compiling interview corpora

NorwC
Norwegian NL1

LINDSEI-NO
English IL

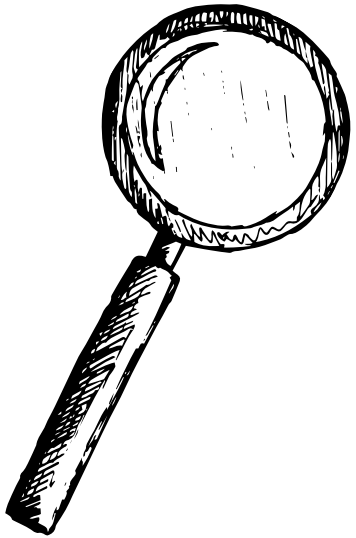
Our research question:

How can a spoken learner corpus be compiled to make valid claims about utterance fluency variations?

LINDSEI: Gilquin, De Cock, & Granger (2010)

INLAND NORWAY UNIVERSITY OF APPLIED SCIENCES

Background:
Zooming in on pause behaviour



- ***Cognitive fluency – utterance fluency – perceived fluency*** (Segalowitz, 2010)
- **"Fluency gaps"** (Segalowitz, 2010) between NL and IL speech (e.g. Ginther et al., 2010; Götz, 2013)

Background:
Some transcription issues

- Turn-taking and segmentation issues:
  - Who “owns” the pause? When does a turn end?

✧ → How can we **identify** pauses reliably and validly?

➢ Smaller data set from two corpora, possible to explore these issues in detail

Background: Transcription conventions

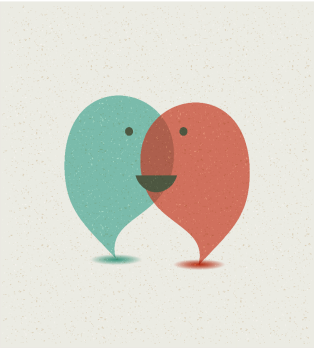
- The LINDSEI project: Minimalistic transcription standard ("broad transcription" (cf. Edwards, 1995))
  - Impressionistic detection of pauses
  - A "linear" representation of speech

- ←→ Specific research needs

"**If not, [the transcript] can hinder detection of patterns of interest, and give rise to directly misleading impressions.**" (Edwards, 1995, p. 19)

Background:
Transcribing pause behaviour in dialogue

- ***Dialogue fluency*** (Tavakoli, 2016)
  - "the decisions researchers make about the measurement of fluency in a dialogue may affect the different temporal aspects of L2 fluency" (p. 147)
- ***Turn pauses*** (Tavakoli, 2016; Peltonen, 2017; Witton-Davies, 2014)
  - Some correlations found between turn pause frequency and length and individual fluency measures: "highlight the importance of including these measures as indicators of fluency when interactional data are examined" (Peltonen, 2017, p. 11)



➢ "approaching fluency solely from the perspective of an individual's (cognitive) competence is not sufficient for characterizing fluency in an interactional setting" (ibid., p. 11)

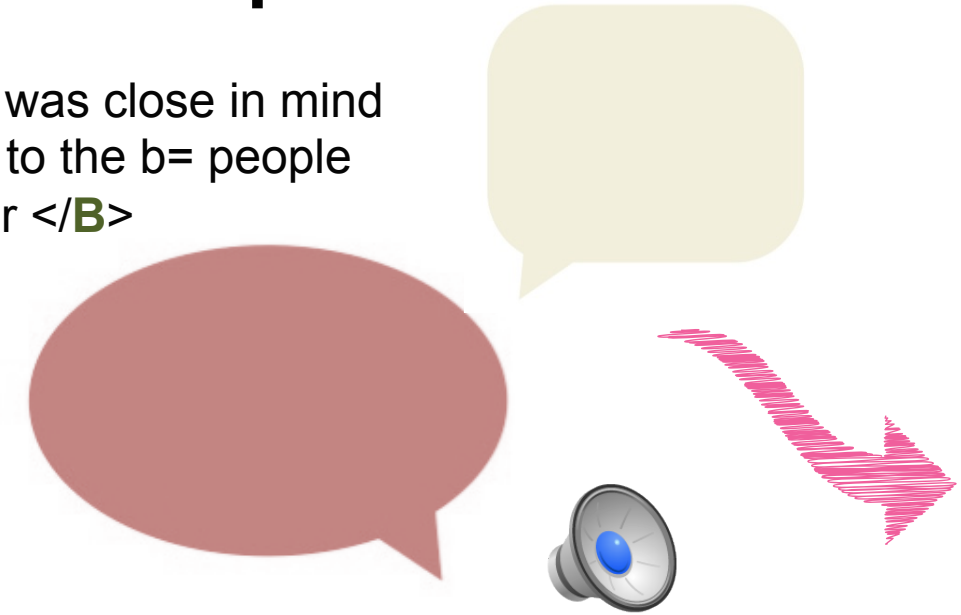
First issue:
Who “owns” the pause?

<B> that's just an experience that was close in mind because I just came from . talking to the b= people about the[i:] (eh) (eh) deadline . for </B>

<A> (mm) </A>

<B> sending in the papers </B>

<A> yeah . okay </A>



LINDSEI-NO, Informant 1

Original LINDSEI-NO transcription

Alternative transcription: Dialogical approach to the segmentation of speech

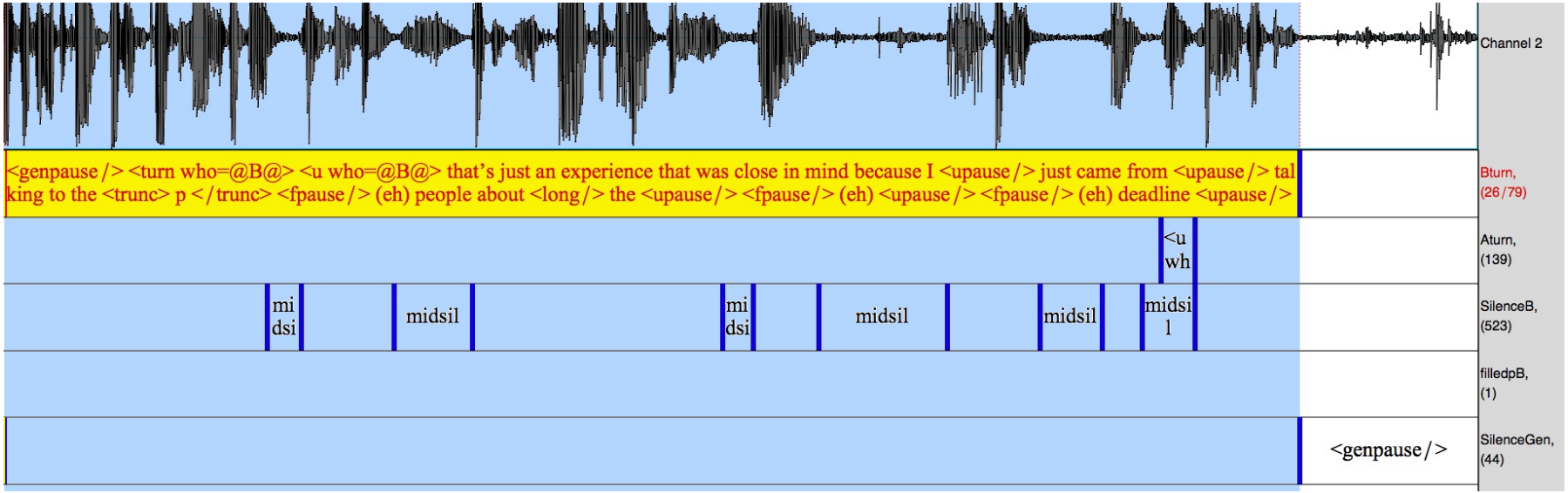
- **Contributing utterance**
  - Utterances from the speaker who holds the turn
- **Non-contributing utterance**
  - Utterances from the speaker who does **not** hold the turn, which do **not** result in turn change

- **Individual pauses**
  - "belonging" to a single speaker
- **Shared pauses**
  - (cf. "turn pauses" (Tavakoli (2016) and Peltonen (2017))

Linell & Gustavsson (1987); Linell (2001)

INN.NO

Alternative transcription: Who owns the pause?



Pauses = silences >0.25 s

(cf. Goldman-Eisler, 1968; De Jong & Bosker, 2013)

PRAAT tool: Boersma & Weenink, 2013
Extraction script: Lennes, 2011

INLAND NORWAY UNIVERSITY OF APPLIED SCIENCES

Alternative transcription: Who owns the pause?

***B turn, B utterance*:**

that's just an experience that was close in mind because I (0.31) just came from (0.73) talking to the p= (eh) people about the: (0.28) (eh) (1.2) (eh) deadline (0.58) for (0.49) *(response)* sending in the papers **(*end B utterance*)**

A utterance: *(response)* (mm) **(*end A utterance*) (*end B turn*)**

[Shared pause, 2.76]

***A turn A utterance*:**

yeah (pause) okay ***(end A utterance) (end A turn)***

LINDSEI-NO, Informant 1

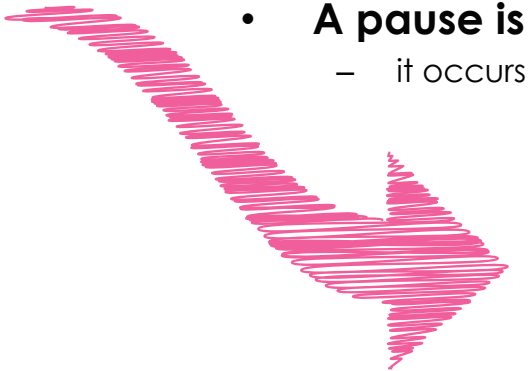
- A's first backchanneling = non-contributing utterance

INN.NO

Summary: Guiding transcription principles

- **A pause is considered part of a speaker's utterance (an individual pause) if it occurs:**
  - turn-initially after a direct question
  - turn-medially or utterance-medially
  - turn-finally, if the turn is viewed as interrupted

- **A pause is considered shared between the speakers in a dialogue if:**
  - it occurs between completed turns (consisting of contributing utterances)



Requires:
A segmentation of speech into turns and (contributing and non-contributing) utterances

INLAND NORWAY UNIVERSITY OF APPLIED SCIENCES

Shared pauses as a possible indication of dialogue fluency and individual fluency?

NL1 (Norwegian)			IL (English)		
5N	2.09	%	3E	1.94	%
3N	3.07	%	1E	2.15	%
1N	4.46	%	4E	2.84	%
2N	4.56	%	5E	3.34	%
4N	4.63	%	2E	6.28	%
6N	5.65	%	6E	6.37	%

Table 1 Shared pause time ratio for each informant's interviews (NL1 and IL), ranked

NL1 (Norwegian)			IL (English)		
3N	14.13	%	1E	18.66	%
1N	20.57	%	5E	19.03	%
5N	20.66	%	3E	19.57	%
4N	21.66	%	2E	24.17	%
6N	21.67	%	4E	31.52	%
2N	23.39	%	6E	29.56	%

Table 2 Individual pause time ratio in each informant's interviews (NL1 and IL), ranked (data from Aas & Rørvik, in press)

Original transcription	Alternative transcription
<B> that's just an experience that was close in mind because I just came from . talking to the p= people about the[i:] (eh) (eh) deadline . for </B> <A> (mm) </A> <B> sending in the papers </B> <A> yeah . okay </A>	*<turn who="B"> <u who="B"> that's just an experience that was close in mind because I* ***<upause dur="0.31"/>*** *just came from* ***<upause dur="0.73"/>*** *talking to the <trunc> p </trunc> <fpause/> (eh) people about <long/> the* ***<upause dur="0.28"/>*** *<fpause/> (eh)* ***<upause dur="1.2"/>*** *<fpause/> (eh) deadline* ***<upause dur="0.58"/>*** *for* ***<upause dur="0.49"/>*** *<response/> sending in the papers </u> <u who="A"> <response/> (mm) </u> </turn>* ***<spause dur="2.76"/>*** *<turn who="A"> <u who="A"> yeah <upause/> okay </u> </turn>*  LINDSEI-NO, Informant 1
• = 2 **B** (individual) pauses	• = 6 **B** (individual) pauses

Second issue: How many pauses?

Informant	AT	OT	Difference
1E	252	123	129
2E	293	311	-18
3E	236	248	-12
4E	328	333	-5
5E	208	242	-34
6E	227	216	11

Table 3 Number of Individual pauses in the alternative transcription (**AT**) and the original transcription (**OT**) of the IL English interviews

→ Utterance fluency vs. (transcriber's) perceived fluency

Second issue: How many pauses?

Informant	AT	OT	Difference
1E	252	123	129
2E	293	311	-18
3E	236	248	-12
4E	328	333	-5
5E	208	242	-34
6E	227	216	11

Table 3 Number of Individual pauses in the alternative transcription (**AT**) and the original transcription (**OT**) of the IL English interviews

Speaker 1 (more pauses in the alternative transcription):

- Pauses in conjunction with non-contributing utterances from the interlocutor
- Pauses in conjunction with other disfluency phenomena (in disfluency clusters)
- Pauses at syntactic junctures
- Silences in conjunction with noticeable breathing
- The transcriber "blinded" by other fluency variables?

Speaker 5 (fewer pauses in the alternative transcription):

- Many short pauses (shorter than the cut-off point of 0.25 s)
- Pauses in overlap situations
- Other hesitation phenomena picked up by the transcriber as pauses?

→ Utterance fluency vs. (transcriber's) perceived fluency

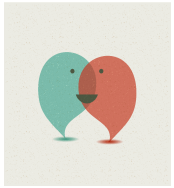
Conclusion

R.Q.: How can a spoken learner corpus be compiled to make valid claims about utterance fluency variations?

- **Issues related to turn-taking** are significant from the perspective of **fluency research.**
  - Turn-initial, turn-medial, and turn-final pauses
- **Dialogue fluency: Utterances should not be viewed as independent of their immediate co-text** .
- **A small set of manageable criteria related to the dynamics of the conversation** can create a more valid picture of pause behaviour, and should also be included in the compilation and transcription/modification of **large-scale spoken corpora.**
- **Lays the ground for** more valid contrastive fluency research, viewing pause behaviour **within and across languages.**

INN.NO

References



- Aas, H. L., & Rørvik, S. (in press). Investigating individual pause profiles through the use of a comparable NL1/IL corpus.
- Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer (Version 5.3.56) [Computer Software]. Available from: http://www.praat.org.
- De Jong, N. H., & Bosker, H. R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In R. Eklund (Ed.), Proceedings of Disfluency in Spontaneous Speech, DiSS 2013 (pp. 17-20). Stockholm: Department of Speech, Music and Hearing.
- Du Bois, J. W., Cumming, S., Schuetze-Coburn, S., & Paolino, D. (Eds.). (1992). *Santa Barbara Papers in Linguistics, Discourse transcription* (Vol. 4).
- Chafe, W. L. (1980). Some reasons for hesitating. In H. W. Dechert & M. Raupach (Eds.), *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler* (pp. 169-180). The Hague: Mouton Publishers.
- Edwards, J. A. (1995). Principles and alternative systems in the transcription, coding and mark-up of spoken discourse. In G. Leech, G. Myers, & J. Thomas (Eds.), *Spoken English on computer. Transcription, mark-up and application* (pp. 19-34). Harlow: Longman Group Limited.
- Gilquin, G. (2008). Combining contrastive and interlanguage analysis to apprehend transfer. Detection, explanation, evaluation. In G. Gilquin, S. Papp, & M. B. Díez-Bedmar (Eds.), *Linking up contrastive and learner corpus research* (pp. 3-33). Amsterdam: Brill | Rodopi.
- Gilquin, G., De Cock, S., & Granger, S. (2010). *Louvain International Database of Spoken English Interlanguage. Handbook and CD-ROM*. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing, 27*(3), 379-399.
- Goldman-Eisler, F. (1968). Psycholinguistics: Experiments in spontaneous speech. London: Academic Press.
- Götz, S. (2013). *Fluency in Native and Nonnative English Speech*. Amsterdam: John Benjamins.
- Lennes, M. (2011). SpeCT - The Speech Corpus Toolkit for Praat. Retrieved from http://www.helsinki.fi/~lennes/praat-scripts/
- Linell, P. (2001). *Approaching Dialogue*. Amsterdam: John Benjamins.
- Linell, P., & Gustavsson, L. (1987). *Initiativ och respons: Om dialogens dynamik, dominans och koherens*. Linköping: Department of Communication Studies.
- McCarthy, M. (2010). Spoken fluency revisited. *English Profile Journal, 1*(1), 1-15.
- Peltonen, P. (2017). Temporal fluency and problem-solving in interaction: An exploratory study of fluency resources in L2 dialogue. *System, 70*, 1-13. doi:10.1016/j.system.2017.08.009
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.
- Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *IRAL, 54*(2), 133-150.
- Witton-Davies, G. (2013). *The study of fluency and its development in monologue and dialogue*. (Doctoral dissertation), Lancaster University, Lancaster.