FILLERS AND PLACEHOLDERS: AN INTRODUCTION
FRANÇOISE ROSE
ATELIER MORPHOSYNTAXE, DDL, LYON, 23 FEB. 2021
OVERVIEW

- A quick introduction to disfluencies
- Placeholders in the literature
- Fillers and placeholders in text analysis

NB: Typologically understudied domain

  Patchwork presentation

  Terminology does not always seem to be standardized

  Different possible approaches (speech processing, interaction, syntax...)

2
A QUICK INTRODUCTION TO DISFLUENCIES
SPEECH DISFLUENCY

- Syntax-for-conversation
- **Disfluencies** = interruptions in the regular flow of speech
  - such as using *uh* and *um*, pausing silently, repeating words, or interruption for self-correction
  - vs. **speech errors** = wrong words or speech sounds, not necessarily with an interruption
- Disfluency is common.
Disfluencies commonly stem from delays or errors in the cognitive processes of language production.

- More frequent before (low-frequency) content words, longer utterances or when the topic is unfamiliar (Corley et al 2008:590); for abstract rather than concrete objects (Reynolds and Pavio 1968).
- But also varies depending on context (choice): contextual probability; humanities vs natural sciences; talking to humans vs machines; in dialogues vs monologues (Corley et al 2008:591).
- It is argued that disfluency can be used with communicative goals (Corley et al 2008:592).
  - It facilitates comprehension in some situations, such as: confidence of the speaker, speaker wants to keep the floor, prediction by the listener that an upcoming referent is difficult for the speaker...
  - = intentional use ‘pay attention, the speaker is in trouble and the next part of the message might not be what you predicted’; or listeners sensitivity to patterns?
Repair: practice that the participants in interaction use to tackle problems they encounter in speaking, hearing, and understanding talk (Schegloff, Jefferson, and Sacks 1977)

Not to be taken as necessarily resulting from an error and involving a replacement:

- Word replacement (i.e. lexical correction)
- Correction in person reference
- Next-speaker selection
- Time for thinking
- Sentence abortion
- Request for recipient-gaze
- Management of overlapping talk
- Production of a dispreferred...
SELF-REPAIRS VS OTHER-REPAIRS (SCHEGLOFF ET AL 1977)

- Self-initiated vs. other-initiated
- Self-achieved vs other-achieved repair
  - Other-initiation overwhelmingly yield self-corrections
  - Preference for self- over other-correction

Examples from Schegloff et al (1977:364, 366)

Ken: Is Al here today?  
Dan: Yeah.  
    (2.0)  
Roger: → He is? hh eh heh  
Dan: → Well he was.
REPAIRS (SCHEGLOFF ET AL 1977)

- Successful or not
- Initiated within the same turn, within the next turn, or the third turn
- Most self-initiated repairs are successful within the turn in which they are initiated vs. other-initiated repairs usually take several turns to be successful

Examples from Schegloff et al (1977:364, 366)

C: C’n you tell me- (1.0) D’y you have any records → of whether you- whether you- who you sent- → Oh(hh) shit.

G: What’d you say?

C: I’m having the worst trouble talking.

Hannah: And he’s going to make his own paintings.
Bea: Mm hm,
Hannah: → And- or I mean his own frames.
Bea: Yeah,
OTHER-INITIATED REPAIRS (SCHEGLOFF ET AL 1977)

- Strategies for other-initiated repairs Schegloff et al (1977:367-368)
  - *Huh? What?*
  - *Who? Where? When?*
  - Partial repeat + question word. Ex: *Met whom?*
  - Partial repeat
  - *Y’mean* + possible interpretation

- Not further discussed in this Atelier
  - See work by Dingemanse et al.
  - Requires conversational data
SAME-TURN SELF-REPAIRS

- **Word search**

  When speakers search for a word, devices for delaying the word:
  - on the current word: phonological devices
    - sound stretch
    - cut-off
  - after current word:
    - silence
    - audible breath pauses (?)
    - **fillers**: all non-silent devices deployed after the current word to delay the next word due
      - Fr. bouche-trou

  Example from Schegloff et al. 1977: 363

  *B’t, a–another one theh wentuh school with me*
  →  *waːs a girl naːmed uh, (0.7) *W’t th’ hell wz*
FILLERS

- Fillers
  - Interjective hesitators
    - non-lexical conventionalized sounds, a.k.a search sounds or filled pauses (Eg. ‘uh’, ‘um’) [debate on whether these are words, see Corley et al 2008]
    - lexicalized items, often from demonstratives (Sp. ‘este’) or interrogative pronouns (Hayashi & Yoon 2010)
    - discourse markers (Eg. ‘y’know’, ‘like’)
  - Explicit editing terms
    - overt statement from the speaker, recognizing the existence of disfluency, e.g. ‘rather’, ‘sorry’
  - Placeholders
    - a lexical substitute for a delayed constituent
    - a.k.a. lexical fillers
    - Fr. joker
FILLERS & FUNCTIONS: DIFFERENT FILLERS

- In English (Clark and Fox Tree 2002)
  - *Uh* signals minor delays while *um* signal major delays

- In Finnish (Sorjonen and Laakso 2005)
  - *eiku* projects that the speaker will replace something she has said, or will abandon the line of talk and syntactic construction she has been building up; [correction]
  - *tai* projects that the speaker will replace something she just said but the element replaced will still remain as an alternative to the new element; [addition]
  - *siis* typically projects that the repair is done in order to specify further or explain something the speaker just said [specification]

- In Danish (Navarretta 2015)
  - *øh* mainly for speech planning, before a content word (lexical retrieval)
  - *øhm* generally for feedback, before clauses, often after conjunctions
  - *mm* in most cases as feedback signal, used alone.
Most work done on English

- Inverse frequency relation between hand gestures and filled pauses (see ref in Navarretta 2015)
  - Also more head/face/body gestures with silent pauses and audible breath pauses than with filled pauses (see ref in Navarretta 2015)
- Fillers often (2/3) accompanied by gestures, especially head gestures (Navarretta 2015, for Danish)
- The functions of the gestures correspond to the function of the filler.
  - Many hold gestures co-occur with filled pauses (see ref in Navarretta 2015), signalling the planning of new spoken content and new gestures (not with fillers used for feedback).
FILLERS & SYNTAX: RECYCLING

- **Recycling strategy**: material preceding the ‘error’ is repeated in the correction or not (see below on placeholders)
- **Recycling scope**: “When English speakers recycle part of an utterance, they do not back up a random number of words [...] speakers returned either to the beginning of a word within the phrasal constituent under construction when repair was initiated or to the beginning of the clause” (Fox, Hayashi and Jasperson 1996: 186)
FILLERS & SYNTAX: RECYCLING (FOX, HAYASHI AND JASPERSON 1996: 204-206)

- English: recycling strategy with scope over the whole constituent (NP, PP or VP) or the clause

  on the back of his pickup truck [with a,*] (0.4) with a jack.

- Japanese (no articles, postpositions). Either no recycling and use of a placeholder (a demonstrative); or constituent-internal recycling. No clause recycling, no full constituent recycling.

  tteyunika koko denwa  [kaket-*] kakete kite sa,
  I.mean here telephone ca- call come FP
  'I mean, (they) ca- called us here,'

- Difference explained by the fact that the utterances have a tighter syntactic organization in English than in Japanese (projection).
**MORPHOLOGICAL REPAIRS**

- **Morphological repairs** (in Kärkkäinen, Sorjonen, and Helasvu 2007)
  - rare?
  - depends on language type?

- Found in Japanese but not in English (Fox, Hayashi & Jasperson 1996: 201-203)
  - « verb endings in English are more tightly "bonded" to the verb than are verb endings in Japanese and hence are less available for individual replacement”
    - Japanese endings are full syllables vs. English consonantal
    - Japanese agglutinating vs. English fusional
    - Japanese adverbial suffixes vs. English subject agreement

---

Japanese (Fox, Hayashi & Jasperson 1996: 202)

```
ja nanji goro ni kirdai(shi-)sso?
```

> Then about what time (shall we) go out?

English (Fox, Hayashi & Jasperson 1996: 202)

```
She [looked]-"s at the table.
```

16
Disfluencies (pauses) as evidence for boundary between proclitic and lexical host (Himmelman 2014)

Filler as an indicator of discourse segment boundaries (Watanabe 2002)

Filler as a tool for morphological analysis (Watanabe 2010)

- Identification of clitics in Sliammon Salish: The filler does not occur within the (morphological) word domain; therefore, if this filler occurs between a morpheme and a word, then that morpheme must be a clitic or another word, not an affix.
PLACEHOLDERS IN THE LITERATURE
MAINLY BASED ON PODLESSKAYA (2010)
A placeholder:

- is a referential expression that is used as a substitute for a specific lexical item that has momentarily eluded the speaker (and which is often specified subsequently as a result of a word search), and...
- it occupies a syntactic slot that would have been occupied by the target word, and thus constitutes a part of the syntactic structure under construction (Hayashi & Yoon 2006:490).
- = a word with minimal lexical content, in many cases carrying the appropriate morphology

Russian

On kupti-l vsjakie et-i... pirozny-e.
he buy-PAST.SG.M various PH-ACC.PL cakes-ACC.PL

"He bought various PH [whatchamacallit] ... cakes."
Lit. "He bought various them ... cakes."
Precategorization of the delayed constituent

- Selection of placeholder (when several are available) anticipates the paradigmatic class of delayed constituent (word class, gender, type of referent...)
- Like indefinite and interrogative pronouns, or classifiers in indefinite expressions.

Examples:

- Eg. *whatchamacallit* and *thingamajig* for objects, *whatsisname* and *whatsername* for persons and depending on gender (Channel 1994: 160, Enfield 2003:105)
- Agreement in gender/class/classifiers
- In Nahavaq, distinct fillers for common nouns, persons and locations (Dimock 2010 in FPP)
- Fr. ‘bidule’ for smaller items than ‘truc’ or ‘machin’
‘Placeholder fillers thus reveal in a most remarkable way the fact that participants attend to syntactic projection, and that when difficulties arise in pursuing that projection, items of just the right morpho-syntactic sort are created to momentary fulfill the syntactic projection.’ (Fox 2010:3 in FPP)

Interactional perspective:
- speaker has additional time for accessing the target word
- speech is not interrupted
- the recipient can be recruited to provide help with finding the target word
Many languages use placeholders instead of nominals only. In other languages, these nominal placeholders also extend their use to verbs and even adjectives and adverbs. Postulated scale N > V > Adj/Adv.

Ulcha (Sumbatova et al. 2005, by courtesy of N. Sumbatova & S. Toldova)

a. xaj-tʃ-n gær-tʃ-n don-do-ro.
   PH-LAT-3SG branch-LAT-3SG alight-IMP FCNV

b. ami-gdolo xaj-xa mapanaŋ-jo-xa.
   father-PTCP PH-PAST get.older-PASTP
   'Father got PH [whatchamacallit] ... got older.'
In other languages, there are some placeholders specific for verbs (verbal placeholders)
  - Often from a verb ‘do’

Uđi
moč:a ěpaža-en-jan ělak: ... he-b-su ... kaci-e.
nettle-DAT special_knife-erg=1pl [press] what-do-pres cut-lv;prs
{From a culinary recipe.} ‘The nettle, with a special knife we press...
wè WHAT-DO, cut’
PLACEHOLDER AND MORPHOSYNTAX

- Morphology on nominal placeholders usually as expected
- Morphology on placeholders used instead of a verb often idiosyncratic
  - Attached on placeholder as it would be on expected constituent, even if prototypically non-verbal
    - Ex: Russian distal demonstrative *eto* with aspect and NEG markers

```
i on pri-eto ... pri-bit.
and he  Pref-PH  Pref-nail.PASTP
‘And he was PREF-PH [whatchamacallit] ... nailed [to ...].’
Lit. “PREF-this PREF-nailed”
```

- Attached to a derived stem. It. *coso* (N) \(\rightarrow\) *cosare* (V), Fr. *machiner*
- On compound made of semantically bleached verbal root ‘do’+ nominal placeholder
  - Possibly idiosyncratic incorporation strategy
PLACEHOLDERS AND MORPHOSYNTAX (REPLICATION/MIRROR)

- Variable replication of the grammatical marking of the delayed constituent
  - Exact mirror
  - Partial mirror
    - Ex. Mojeño Trinitario manje’e takes enclitics, but not prefixes
  - No mirroring: default grammatical form (often least marked form).
    - Ex Russian eto
  - Variable
    - Ex: Teko (Rose 2011: 176)

Mojeño Trinitario
manje’e=ripo na-nok=ripo
um=PFV 3PL-put=PFV
‘They um.... they put it’ 36.12

Russian
Vy nas povedjo-te v eto/eto...
you us take-pres.2PL to PH.accc.sg.m/PH.nom.sg.n
v trenajor-yj zal?
to fitness-acc.sg.m hall.acc.sg.m
‘Will you take us to PH [whatchamacallit] ... to a fitness hall?’

Teko
kob pitan-am ba?e kito-r-ehe e-iba.
cop enfant-transl chose grenouille-reln-postp 3.II-animal.domestique
Il y a un enfant et sa euh... grenouille apprivoisée.
kito-r-a?ir o-iun? ba?e-pope-dji, bokal-a-pe-dji o-iun.25
Il met la petite grenouille dans un truc, il la met dans un bocal.
Variable repetition of preceding material (when placeholder not at the left margin of constituent), a.k.a. « scope of backing up »

- **Recycling strategy**
- Non-recycling strategy

Related to basic patterns of recycling observed in speech disfluencies other that placeholders, e.g. in self-repairs.

- sensitive to general morphosyntactic characteristics of the language, including word-order and phrase structure

- Issue independent from mirroring (Russian example)
PLACEHOLDERS AND DIACRONY

- Sources for placeholders
  - Pronouns: 3rd person, demonstrative, interrogative (in Udi and Agul), indefinite or emphatic (Sp este, Rus. eto)
  - Semantically beached N ‘thing’ (Teko baʔe)
  - Both DEM + N (Lao qan°-nan⁴ ‘that thing’ in Enfield 2003)
  - V ‘do’
  - Clauses, often with interrogative word + naming N or V (Eg. watchamacallit, Kuwaiti Arabian šisma ‘what-name-his’), cf. Enfield 2003
Through bleaching, and/or pragmaticization, lexicalization


- With formal changes, ex: Italian coso (masc. N developed from fem. N cosa ‘thing’)

- Lexical flexibility. Ex. Lao qan⁰-nan⁴ ‘that thing’ used as predicate (Enfield 2003:108)

```
  mê¹  caw⁴  ŋang²  bô⁰  qan⁰-nan⁴  juu¹  tii⁴
mother  2sg  still  not  thing-that  pcl  pcl
‘Your mother’s not yet ‘that thing’, right?’
```

- Default placeholders often tend to develop into interjective hesitation markers undergoing ‘pragmaticization’, and are no longer integrated into the syntactic structure of the clause (Hayashi & Yoon 2006, 2010 in FPP).
PLACEHOLDERS AND SPEECH

- Frequent in speech
  - 5 / 1000 words in Russian narratives (Podlesskaya & Kibrik 2006)
  - 6.68 / 1000 words in Chinese conversations (Zhao & Jurafksy 2005)
  - ~10% of repairs in various languages (Podlesskaya 2010)
Some placeholders are at the heart of the semantics/pragmatic interface (Enfield 2003)

- Demonstratives used without any antecedent or contextually present objects
- Speaker’s assumption about the listener’s knowledge (recognitional deixis)

  \textit{WHAT-d’you-call-it:}
  - Something
  - I can’t say the word for this thing now
  - By saying \textit{WHAT-d’you-call-it} I think you’ll know what I’m thinking of.

- Additional functions
  - Avoidance/conspiracy: not « I can’t say the word » but « I don’t want to say the word »
  \textit{Did you bring any you-know-\textit{WHAT}?} (e.g. ‘marijuana’, to a party)
  \textit{I saw John at the club again with you-know-\textit{WHO}.}
  - Euphemistic use, to soften face-threatening acts. Ex: use of \textit{qan\textsuperscript{0}-nan\textsuperscript{4}} ‘that thing’ in Lao.
PLACEHOLDERS/FILLERS AND SPEECH

- Interjexional function
  - Interaction management
    - Signaling beginning of turn
    - Signaling that the speaker is thinking to take the floor in conversation
    - Feedback giving
    - Feedback eliciting
  - Cognitive processes of discourse planning
    - to alert the recipient to a new topic or a name
    - to mark sensitivity of the action by postponing the crucial part, displaying lexical access problems and eliciting assistance for that...
PLACEHOLDERS AND SPEECH

- **Intentional use** = not a non-deliberate hesitation, but a controlled rhetorical device and a cue for the recipients.
  - Interjectional function
  - As vague or imprecise expressions
As vague or imprecise expressions (when exact name is unnecessary or inappropriate):

- instead of the exact correlate = generic expressions. Cf. Fr. *truc* (see also Amiridze 2010 in FPP with verbs in Georgian)
- together with the exact correlate, to indicate lower responsibility for the nomination = vague category identifier ("N or whatever, une sorte de N").
- Grammaticalization towards similative/associative.
- Functional proximity (and possibly shared source) between placeholders and cataphoric pronouns

Armenian (Podlesskaya & Xurşudjan 2006)

\[
\begin{align*}
\text{hac}^h & \text{-ic}^h & \text{ban}^- & \text{-ic}^h & \text{her}.
\text{bread-ABL} & \text{PH-ABL} & \text{bring.IMP.2SG}
\text{‘Bring [us] bread and all that stuff!’ Lit. ‘bread PH bring.’}
\end{align*}
\]
FILLERS & PAUSES IN TEXT ANALYSIS
FILLERS & PAUSES IN TEXT ANALYSIS

- Transcription conventions
- Partage de pratiques sur les bases de données textuelles (DDL)
- Fox, Hayashi & Jasperson (1996)’s coding
False starts, repetitions, and hesitations (‘uhm’ and the like) are usually edited out. ”

Whatever mistakes or inconsistencies have been included in the transcript will be carried on to these other levels of analysis. » (Himmelmann 2006:259 The challenges of segmenting spoken language)
TRANSCRIPTION CONVENTIONS

- Often from conversation analysis
  - Compatibility with software for morphosyntactic analysis?
- Others to be shared?
PARTAGE DE PRATIQUES SUR LES BASES DE DONNÉES TEXTUELLES

- Séances en 2014
- Thèmes abordés pertinents
  - Pauses
  - Mots tronqués
  - Hésitations
  - Faux départs
- Pratiques de plusieurs membres du labo
  - Copies d’écran
  - Evaluation générale: 😊 😊
  - Analyse des avantages et inconvénients
- Archivés et accessibles à:
  - Tsetse (pérenne): tsetse.dl.cnrs.fr/DDL/Ateliers/Pratique-BDD-textuelles
  - Google Drive (non-pérenne): https://drive.google.com/drive/folders/0BwMLJAG1WwMjQ1BrbG1oSWJabTg?usp=sharing
ROSE - PAUSES
ROSE - PAUSES

Avantages:
- Pauses indiquées de manière constante
- Pas besoin de quantifier la pause
- Glosable (pas de symbole d'erreur)
- Très utile pour l'analyse

Inconvénients:
- Pas conventionnel
- À modifier pour exemples dans articles ou thèse
- Quand rupture syntaxique claire, usage de virgule ou point
ROSE – FAUX DÉPARTS
ROSE – HÉSITATIONS
Data coding: repairs were coded for

- (1) the phrasal syntactic constituents, including clause, that were under construction when the repair was initiated (e.g., noun phrase and clause or verb phrase and clause)
- (2) the syntactic organization of the repair (e.g. just the repairable was repeated, or the entire clause was redone)
FURTHER TALKS AT THE ATELIER
FURTHER TALKS

- Further dates
  - 6 avril 2021
  - 11 mai 2021
  - 8 juin 2021
  - 22 juin 2021
- Planned for either two short talks or one long talk
- More dates could be added is necessary
BRAINSTORMING

- Study of disfluency in a corpus
- Sketch of fillers in a language / language family
  - With possible focus on one item
  - Looking at placement and function in discourse
- Overview on « truc, machin, bidule » in French (short reference list available) – or guest ?
- Fillers and pauses
  - DoReCo
- Fillers and gestures
  - Video corpora

- Overview of literature on transcription conventions – or guest ?
- Sharing practices in transcription (update)
- Interest for Dendy ?
  - In acquisition
  - Oral vs written disfluencies
  - Pathologies (see paper by Davis and Maclagan 2010 on Alzheimer in FPP)
- ...

...
SUGGESTIONS

- Denis Bertet on Tikuna
- Françoise Rose on Teko and Mojeño Trinitario
- Christian Döhler on Komnzo and Bine (New Guinea)?
All chapters of following book (abbreviated as FPP)

- Especially the typological chapter:
OTHER REFERENCES


