The Disfluency, Exclamation, and Laughter in Dialogue (DUEL) Project

Jonathan Ginzburg , Ye Tian,
Pascal Amsili, Claire Beyssade
Barbara Hemforth, Yannick Mathieu
Claire Saillard
CLILLAC-ARP (EA 3967) &
Laboratoire Linguistique Formelle (LLF) (UMR 7110)
& Laboratoire d'Excellence (LabEx)—EFL, Université Paris-Diderot, Paris, France

Julian Hough, Spyros Kousidis,
David Schlangen
Faculty of Linguistics and Literary Studies,
Bielefeld University, P.O. Box 10 01 31, 33501,
Bielefeld, Germany

`yonatan.ginzburg@univ-paris-diderot.fr`
`http://www.dsg-bielefeld.de/DUEL/`

Abstract

The aim of the 3-year Disfluency, Exclamation, and Laughter in Dialogue (DUEL) project between Université Paris Diderot (Paris 7) and Bielefeld University is to model the human capacity for speaking and understanding disfluent and laughterful utterances, and to create formal models and computational systems capable of this processing. The other challenge in this enterprise is to model this interaction incrementally, that is, online as it happens word-by-word in real dialogue.

1 Introduction

Although disfluencies, exclamations and laughter occur frequently in spoken conversation, they have received little attention both within formal theories of grammar, where they are widely perceived as phenomena outside of its range, and practical dialogue modelling, where they are perceived as distractions to be filtered out. The Disfluency, Exclamation, and Laughter in Dialogue (DUEL) project, based at Université Paris-Diderot (Paris 7) and Bielefeld University aims to address this situation by an integrated empirical, theoretical, and computational research programme. The project is funded by the Agence Nationale de Recherche (ANR) and by the Deutsche ForschungGemeinschaft (DFG) within the *projets franco-allemand en sciences humaines et sociales*.

In the rest of this project description, we provide some motivation for the project and describe some of its objectives.

2 Motivation

2.1 Disfluencies in the grammar

Disfluencies are highly frequent in natural language production. They include editing terms such as *uh* and *I mean* as well as repeats—often referred to as *recycling* ('I - uh - I wouldn't', e.g. (Clark and Wasow, 1998)) and revisions. In spoken language, disfluencies are typically found in about six out of 100 words (Fox Tree, 1995) / more than 35% of all utterances (Jurafsky and Martin, 2009, p. 453).

Despite their ubiquitous nature, grammarians have, with very few exceptions, regarded disfluencies as elements not fit to populate the grammatical domain. Their very existence is a significant motivation for the competence/performance distinction (Chomsky, 1965,). (Ginzburg et al., 2014) argue that far from constituting meaningless "noise", disfluencies participate in semantic and pragmatic processes such as anaphora, conversational implicature, and discourse particles, as illustrated in (1):

(1) a. Peter was, well, he was fired. (Example from (Heeman and Allen, 1999); anaphor refers to material in reparandum.)

    b. A: Because I, any, anyone, any friend, anyone, I give my number to is welcome to call me (Example from the Switchboard corpus (Godfrey et al., 1992,); implicature based on contrast between repair and reparandum: *It's not just her friends that are welcome to call her when A gives them her number.*)

    c. The other one did, no, other ones did it. (Example from BNC (file KB8, line 1705); material negated by *no* originates in the reparandum.)

Beyond this, (Ginzburg et al., 2014) offer detailed argumentation for why disfluencies do belong in the grammar. In particular, they point out that disfluencies exhibit linguistic regularities across all levels of grammatical representation, cross-linguistic variation, and universals. All these

are hallmarks of processes that need representation in a grammar. Crosslinguistic variation has been documented in some detail in comparative work between morphosyntactic aspects of repair on a wide range of languages by Fox and collaborators (e.g., (Fox et al., 1996; Wouk et al., 2009; Fox et al., 2010)) and in and in phonetic analysis of hesitation markers (Candea et al., 2005,).

Understanding the range of cross-linguistic variation and the scope of universals in the area of disfluency is one of the motivations for the cross-linguistic programme of DUEL, where a parallel corpus in French, German, and Chinese will be compiled.

2.2 Laughter in the grammar

Laughter is multifunctional (Glenn, 2003). (Schegloff, 2001) illustrates the force cancelling effect of laughter (e.g. in indicating an utterance is not to be taken seriously or in enabling a socially delicate utterance to be made without causing offence.):

(2) Freda: Becaus-ah
(silence: 3.3 seconds)
Rubin: They don mind honey they're jus not gonna talk to us ever again.
Dave: =(laugh: hehem)/(ri: (h)ight)
Kathy: We don mind, we jus ne:ver gonna talk to you e:__ver (laugh) hh(h'g)
Dave: No, b't
Rubin: (laugh) heheheheh

Laughter in its intra-sentential occurrence bears a strong relation to disfluency in enabling a speaker to express uncertainty about the force of the utterance they are making:

(3) A: [I,+I] [d,+ don't] feel comfortable about leaving my kids in a big day care center,
B: Worried that they're not going to get enough attention?
A: Yeah, and uh you know colds and things like that [laughter] (From Switchboard)

3 Objectives

3.1 Experimental Work

Interaction will be recorded in French (Paris), German (Bielefeld), and Chinese (both sites). This ensures variability both with respect to possible morphological and syntactic constraints on the placement of the phenomena of interest as well as to possible cultural differences in their discourse use. Chinese is chosen since its morphological properties lead us to expect significant variation with respect to disfluencies in polysyllabic and inflectionally rich French and German; conversely, given that the basic SVO word order of Chinese resembles French quite a bit more than German, this will also enable to control for the role of word order v. morphology.

3.2 Theoretical Work

The goal of this work area is to extend work on grounding, clarification interaction, and disfluency within the framework of KoS (Ginzburg, 2012; Ginzburg et al., 2014) l so that it can both underpin the analysis of various linguistic phenomena revolving around disfluencies, laughter, and interjections, serve as the grammar and dialogue theoretical basis for computational work, based on the parallel corpus we will have compiled. There are two main formal tasks in this area: first, develop KOS_{incr}, a detailed, principled incremental semantics for dialogue in KoS, using Type Theory with Records (Cooper, 2012). Second, expand KOS_{incr} to KOS_{incr}^{EMA}, a dialogical theory whose states encode emotive appraisal (Marsella and Gratch, 2009). Each of these formal innovations will underpin detailed linguistic analysis.

3.3 Computational Work

The goal of this work area is to provide a practical, computational model of disfluency and laughter in dialogue, which captures the subtleties observed in the data and implements the main elements of the theory. The model will be implemented within an (extension of an) existing dialogue system (Buß et al., 2010), and will be evaluated for the improvements it effects in perceived naturalness of the behaviour of the system.

We aim to build a system that can be disfluent in a natural way, and is also capable of interactionally appropriate laughter when interacting with users. These are milestones for moving towards more natural spoken conversations between humans and machines.

References

Okko Buß, Timo Baumann, and David Schlangen. 2010. Collaborating on utterances with a spoken dialogue system using an isu-based approach to incremental dialogue management. In *Proceedings of the SIGdial 2010 Conference*, pages 233–236, Tokyo, Japan, September.

M. Candea, I. Vasilescu, M. Adda-Decker, et al. 2005. Inter-and intra-language acoustic analysis of autonomous fillers. In *Proceedings of DISS 05, Disfluency in Spontaneous Speech Workshop*, pages 47–52.

N. Chomsky. 1965. *Aspects of the Theory of Syntax*. MIT Press.

H.H. Clark and T. Wasow. 1998. Repeating words in spontaneous speech. *Cognitive psychology*, 37(3):201–242.

Robin Cooper. 2012. Type theory and semantics in flux. In Ruth Kempson, Nicholas Asher, and Tim Fernando, editors, *Handbook of the Philosophy of Science*, volume 14: Philosophy of Linguistics. Elsevier, Amsterdam.

B.A. Fox, M. Hayashi, and R. Jasperson. 1996. Resources and repair: A cross-linguistic study of syntax and repair. *STUDIES IN INTERACTIONAL SOCIOLINGUISTICS*, 13:185–237.

Barbara A Fox, Yael Maschler, and Susanne Uhmann. 2010. A cross-linguistic study of self-repair: Evidence from english, german, and hebrew. *Journal of Pragmatics*, 42(9):2487–2505.

Janet E. Fox Tree. 1995. The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, 34:709–738.

Jonathan Ginzburg, Raquel Fernández, and David Schlangen. 2014. Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics*, 7(9):1–64.

Jonathan Ginzburg. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford University Press, Oxford.

Phillip J Glenn. 2003. *Laughter in interaction*. Cambridge University Press Cambridge.

John J. Godfrey, E. C. Holliman, and J. McDaniel. 1992. Switchboard: Telephone speech corpus for research and devlopment. In *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing*, pages 517–520, San Francisco, USA, March.

Peter A. Heeman and James F. Allen. 1999. Speech repairs, intonational phrases and discourse markers: Modeling speakers' utterances in spoken dialogue. *Computational Linguistics*, 25(4):527–571.

Daniel Jurafsky and James H. Martin. 2009. *Speech and Language Processing*. Prentice Hall, New Jersey, 2 edition.

S.C. Marsella and J. Gratch. 2009. Ema: A process model of appraisal dynamics. *Cognitive Systems Research*, 10(1):70–90.

E.A. Schegloff. 2001. Getting serious: Joke → serious' no'. *Journal of pragmatics*, 33(12):1947–1955.

F Wouk, B Fox, Makoto Hayashi, Steven Fincke, Liang Tao, M Sorjonen, Minna Laakso, and Wilfrido Flores-Hernandez. 2009. A cross-linguistic investigation of the site of initiation of same turn self repair. *Conversation analysis: Comparative perspectives*.