"Proceedings of the International Conference “Fluency & Disfluency Across Languages and Language Varieties”"

Degand, Liesbeth

Abstract

Within the frame of a research project entitled “Fluency and disfluency markers. A multimodal contrastive perspective”, the universities of Louvain and Namur have been involved in a large-scale usage-based study of (dis)fluency markers in spoken French, L1 and L2 English, and French Belgian Sign Language (LSFB), with a focus on variation according to language, speaker and genre. To close this five-year research project, the (DIS)FLUENCY 2017 International Conference is held in Louvain-la-Neuve on the subject of fluency and disfluency across languages and language varieties. This volume contains the papers presented at the conference.

Document type : *Rapport (Report)*

[p]lin

UCL Université catholique de Louvain

Proceedings of the International Conference

"Fluency & Disfluency Across Languages and Language Varieties"

Université catholique de Louvain

February 15-17, 2017

Louvain-la-Neuve, Belgium

Preface

Fluency and disfluency have attracted a great deal of attention in different areas of linguistics such as language acquisition or psycholinguistics. They have been investigated through a wide range of methodological and theoretical frameworks, including corpus linguistics, experimental pragmatics, perception studies and natural language processing, with applications in the domains of language learning, teaching and testing, human/machine communication and business communication.

Spoken and signed languages are produced and comprehended online, with typically very little time to plan ahead. As a result, they are often characterized by features such as (filled and unfilled) pauses, discourse markers, repeats and self-repairs, which can be said to reflect on-going mechanisms of processing and monitoring. The role of these items is ambivalent, as they can be both a symptom of encoding difficulties and a sign that the speaker is trying to help the hearer decode the message. They should thus be interpreted in context to identify their contribution to fluency and/or disfluency, which can be viewed as two faces of the same phenomenon.

Within the frame of a research project entitled "Fluency and disfluency markers. A multimodal contrastive perspective", the universities of Louvain and Namur have been involved in a large-scale usage-based study of (dis)fluency markers in spoken French, L1 and L2 English, and French Belgian Sign Language (LSFB), with a focus on variation according to language, speaker and genre. To close this five-year research project, the (DIS)FLUENCY 2017 International Conference is held in Louvain-la-Neuve on the subject of fluency and disfluency across languages and language varieties. This volume contains the papers presented at the conference.

This conference benefits from the support of the ARC-project "A Multi-Modal Approach to Fluency and Disfluency Markers" granted by the Fédération Wallonie-Bruxelles (grant nr.12/17-044).

Liesbeth Degand,

on behalf of the organizing committee:

Cédrick Fairon (UCL)
Gaëtanelle Gilquin (UCL)
Sylviane Granger (UCL)
Laurence Meurant (UNamur)
Anne Catherine Simon (UCL)
George Christodoulides (UCL)
Ludivine Crible (UCL)
Amandine Dumont (UCL)
Iulia Grosman (UCL)
Ingrid Notarrigo (UNamur)
Lucie Rousier Vercruyssen (UCL)

Table of content

Keynote speakers

What do listeners think of disfluencies?

Martin CORLEY

University of Edinburg

Recent research shows that listeners are sensitive to disfluencies in unfolding discourse, using them to modify predictions about what will be said. But how do they interpret the pragmatic content of the message? I present a series of experiments examining these issues using a "lying game" in which speakers identify the locations of treasure, either fluently or disfluently. Results show that listeners robustly interpret disfluency as a clue to dishonesty, and that this happens early in comprehension, even where other evidence is available. However, when there is a plausible exogenous cause of the disfluency such as speaker distraction, this effect is modulated, suggesting that listeners are causally interpreting the disfluencies uttered. It is therefore surprising that an interactive version of the paradigm where speakers' utterances are freely generated suggests that listeners' judgements may be misguided: Speakers tend to be more disfluent when telling the truth, although listeners steadfastly continue to distrust disfluent speech.

Do (non-linguistic) variables affect learners' (dis)fluency? A learner corpus-based perspective

Sandra GÖTZ

Justus Liebig University Giessen

It has been noted by various scholars that second-language acquisition (SLA) theory needs to take into account learning context as a determining factor in a learners' spoken fluency (e.g. Norris and Ortega 2001; Freed et al. 2004). Researchers in the field of language testing, as well, have documented a considerable degree of variability in learners' fluency, depending on non-linguistic context variables, such as the test-taker's or interlocutor's gender, the communicative style during the test or the task type (e.g. Porter 1991; Kormos 1999; O'Sullivan 2000; Csépes 2009). However, we face a number of challenges in systematising the multitude of variables that affect learners' fluency performance, because it is very difficult to operationalize them quantitatively. Consequently, empirical research into interrelations between these non-linguistic learning context variables and language learners' fluency far has mainly either focused on a small number of learners (in the fields of SLA and language testing) or a small set of non-linguistic learning context variables (in learner corpus research). Thus, if we want to abstract from the fluency development of individual learners to the general picture of the output of a greater number of typical learners regardless of the learners' L1, we need to take into account comparable data from many learners that have in common a defined set of typical and representative (non-linguistic) variables. This is one of the central rationales in compiling learner corpora as repositories of learners' natural language use (cf. Ellis 1994; Granger 2002).

In this talk, I will discuss how learner corpus research into fluency can benefit immensely from taking into consideration learning context variables gathered from the learner profiles of the *Louvain International Database of Spoken English Interlanguage* (LINDSEI; cf. Gilquin et al. 2010). LINDSEI includes data of learners with 11 different L1s and provides a variety of meta-data on each individual learner in the corpus (e.g. sociobiographic data, languages the learners have been exposed to, the number of years of English instruction at school and at

university, time spent abroad, or other language(s) spoken by the learner), as well as on the interviewers (e.g. gender, languages spoken, familiarity with the interviewee). By way of presenting the findings of some case studies conducted on LINDSEI, I will discuss the effect of these non-linguistic variables on learners' fluency, including the fluency development after a stay abroad or the effect of sociolinguistic variables on the learner's fluency performance regardless of their L1. Finally, the implications of the findings of these case studies will be outlined, particularly focusing on the implications on learner corpus-based fluency research and the fluency development of advanced learners of English.

Csépes, I. 2009. *Measuring oral proficiency through paired-task performance*. Frankfurt am Main: Peter Lang.

Ellis, R. 1994. *The study of second language acquisition*. Oxford: Oxford University Press.

Freed, B., Segalowitz, N. and Dewy, D. P. 2004. "Context of learning and second language fluency in French: comparing regular classroom, study abroad, and intensive domestic immersion programs". *Studies in Second Language Acquisition* 26(2), 275-301.

Gilquin, G., De Cock, S. and Granger, S. 2010. *Louvain International Database of Spoken English Interlanguage*. Handbook and CD-ROM. Louvain-la-Neuve: Presses universitaires de Louvain.

Granger, S. 2002. "A bird's eye view of learner corpus research". In S. Granger, J. Hung and S. Petch-Tyson (eds) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins, 3-33.

Kormos, J. 1999. Simulating conversations in oral-proficiency assessment: A conversation analysis of role plays and non-scripted interviews in language exams. *Language Testing* 16(2):163-188.

Norris, J.M. and Ortega, L. 2001. "Does type of instruction make a difference? Substantive findings from a meta-analytic review", *Language Learning* 51 (s1), 157-213.

O'Sullivan, B. 2000. Exploring gender and oral proficiency interview performance. *System* 28(3), 373-386.

Porter, D. 1991. Affective factors in language testing. In J.C. Alderson & B. North (eds), *Language testing in the 1990s*. London: Macmillan, 32-40.

Modeling disfluencies across domains

Helena MONIZ

Universidade de Lisboa

Disfluencies are on-line editing strategies with several (para)linguistic functions. Everyday we are annalists of our own speech and of others, monitoring distinct linguistic and paralinguistic factors in our communications, using disfluencies to make speech a more error-free system, a more edited message, and a more structured system with coherent and cohesive mechanisms.

This presentation focuses on the analysis of disfluencies, aiming at a characterization of the regular patterns in their production in European Portuguese, and at contributing towards the fully automatic processing of structural metadata events. This analysis was strongly supported on prosodic feature processing, and involved corpora of very different characteristics. For that purpose a framework was built for metadata annotation including prosodic features, a crucial step for Portuguese, since prior to this work our in-house speech recognizer had no integration of such features. This framework allowed us to access several layers of linguistic information (e.g., acoustic-prosodic, POS, pragmatic) in a very flexible way and proved to be a suitable tool for the analysis of metadata events.

The robustness of acoustic-prosodic features across domains was investigated using university lectures and dialogues. Different models trained with one corpus were tested on the

other, revealing that models can be quite robust across corpora for this task, despite their distinct nature. The model trained on dialogues proved to be the more robust one, possibly due to the fact that dialogues contain more contrastive tempo characteristics, while sharing with university lectures most of the pitch and energy patterns on disfluent sequences. Therefore, a model created with such data generalizes better.

In our current research, we try to extend this study to other domains, including human-computer interactions both with virtual and embodied agents.

Signs of (dis)fluency throughout development: The language use of Deaf children who are native users of a signed language considering adult examples of (dis)fluency

David QUINTO-POZOS

University of Texas at Austin

Native users of signed languages are notably perceptive about a stranger's signed language skills. Lifelong signers can typically determine whether someone they have never met before is a native signer, like them, or if the stranger was exposed to the signed language sometime after birth—especially if their exposure did not occur until late in childhood or beyond. Not unlike the perceptual abilities of native users of spoken languages, signers may be utilizing multiple cues—some of which are subtle and others more overt—that signal whether a person had early exposure to the visual-gestural language. Such cues might be considered among the metrics for an individual's *fluency* in a signed language. However, what can be said about judgments of a native speaker's fluency in a signed language, especially if that speaker is a child?

Empirical studies of so-called fluency markers in a signed language are not numerous, although several studies of *fluencemes* in French Belgian Sign Language (LSFB) have highlighted markers of fluency in adult native signers, including rate of signing (Notarrigo & Meurant 2015) and use of repetition (Notarrigo, Meurant, & Simon 2016), among others. Such studies have pushed the boundaries of where signed language researchers look for effects of early language exposure on language processing and use.

Other work on fluency in signed language has questioned the role of multilingualism. For example, a fluent signer of one sign language who also knows another sign language might exhibit examples of a signed "accent". This has been shown for adult signers of American Sign Language (ASL) and Mexican Sign Language (LSM) by Quinto-Pozos (2002, 2008) and signers of Al-Sayyid Bedouin Sign Language (ABSL) and Israeli Sign Language (ISL) by Sandler (2014). These authors share a focus on aspects of phonetics and phonology in the language use of late-learners, highlighting features such as handshape, movement, and so-called 'hand prominence' (fingertip, radial, or ulnar prominence in articulation).

What about fluency within a deaf user's first language, especially during development? Judging a native signer's fluency at various stages in their childhood is particularly challenging task because children are notably variable in their linguistic and psychosocial development. What are the cues that lead a fluent signer to judge a child's signed language use as (less-than) fluent? Are there differences between disfluency that is caused by late exposure to language versus disfluency that co-occurs with a developmental disorder or deficit (e.g., of language and/or cognition)?

I will discuss various aspects of fluency in signed language use. Studies of adults that focus on aspects of phonetics and phonology will be reviewed in order to establish what has been

discussed in terms of second-language (L2) use. In addition, the presentation will highlight aspects of childhood development that may signal unexpected disfluency, according to reports from educators and developmental specialists at schools for the deaf and case study data that we have collected in our lab.

Signed language use by deaf children who are native signers provides a fertile ground for investigating aspects of fluency in the visual-gestural modality throughout development.

Notarrigo, I & Meurant, L. (2015, July). Markers of (dis)fluency across signers' profiles in French Belgian Sign Language (LSFB) A comparative analysis between Native, Near-Native and Late Signers. Poster Presentation. 2nd International Conference on Sign Language Acquisition (ICSLA). Amsterdam, The Netherlands.

Notarrigo, I., Meurant, L., & Simon, A.C. (2016, January). Repetition of signs according to language background in French Belgian Sign Language (LSFB): A comparative analysis between Native, Near-Native and Late Signers. Poster Presentation. Theoretical Issues in Sign Language Research (TISLR) 12. Melbourne, Australia.

Quinto-Pozos, D. (2002). Contact between Mexican Sign Language and American Sign Language in Two Texas Border Town. Unpublished doctoral dissertation. University of Texas at Austin.

Quinto-Pozos, D. (2008). Sign language contact & interference: ASL & LSM. *Language in Society 37*, 2. 161-189.

Sandler, W. (2014). The emergence of the phonetic and phonological features in sign language. *Nordlyd* 41.2: 183-212.

Papers

A contrastive analysis of disfluency markers in Hungarian in four different settings

Mária BAKTI[1], Judit BÓNA[2]

University of Szeged[1], Eötvös University[2]

Speech production in itself is a complex online process, with concurrent phases (Levelt 1989). During consecutive interpreting, a target language text is produced based on the source language text and notes taken based on the source language text; during sight translation the target language text is produced based on a written source language text. Sight translation (ST) is at the boundary of translation and interpreting (Agrifoglio 2004), but some scholars see it as a specific type of written translation or a variant of oral interpretation (Lambert 2004). During ST the interpreter has continuous access to the information in the SL text; during the perception stage, the interpreter must divide attention between visual input and oral production and during text production, the interpreter must also monitor their production while reading and translating.

These factors can determine or influence the process of speech production and speech planning, and thus the disfluency surface markers in the target texts (Bóna and Bakti 2014, Mead 2000, 2002, Shreve et al. 2011)

In this project we examined the output of ten MA students of interpreting in four speech production settings. The ten MA students finished their four-semester interpreter training at a Hungarian university and were preparing for their final exam at the time the recordings were made. Recordings were made in four settings: (1) sight translation from English into Hungarian on the topic of English as a Lingua Franca, (2) consecutive interpreting with notetaking from English into Hungarian on English as a Lingua Franca and its implications for the interpreting profession. Students were asked to prepare a short presentation (3) in Hungarian on English as a Lingua Franca and an informal interview (4) was also carried out with the students in Hungarian in which they spoke about the interpreting tasks and various other topics.

The Hungarian target texts were recorded, transcribed and analyzed for disfluency surface markers.

We aimed to answer the following research questions:

1. What is the pattern of disfluency markers in the four settings?
2. Are there any similarities between the patterns for each setting?
3. What is the function of disfluencies in each setting?

How does timing depend on the setting? (tempo, pauses, other disfluency markers) ?

Our preliminary results indicate that depending on the complexity of the task, the frequency and type of disfluencies changed, in other words the speech production setting does influence speech tempo and the duration and frequency of pauses, together with the occurrence of disfluencies. The most complex production task, sight translation, is characterized by the slowest speech tempo, the longest pauses and longest filled pauses.

Agrifoglio, M. 2004. Sight translation and interpreting. A comparative analysis of constraints and failures. *Interpreting* Vol. 6. No. 1. 43−67.

Bóna, J. és Bakti M. 2014. A beszédtervezés- és kivitelezés temporális sajátosságai szinkrontolmácsok beszédprodukcióiban. [Temporal characteristics of speech planning and execution in the speech production of simultaneous interpreters.] *Fordítástudomány*. XVI. évfolyam, 1. szám. 16-28.

Lambert, S. 2004. Shared Attention during Sight Translation, Sight Interpretation and Simultaneous Interpretation. *Meta* 49(2): 294-306.

Levelt, W.J.M. 1989. *Speaking. From Intention to Articulation*. Cambridge, MA: MIT Press.

Mead, P. 2000. Control of pauses by trainee interpreters in their A and B languages. *The Interpreter's Newsletter* Vol. 10. 89–102.

Mead, P. 2002. Exploring hesitation in consecutive interpreting: An empirical study. In Garzone, G. – Viezzi, M. (ed.): *Interpreting in the 21st century. Challenges and Opportunities*. Amsterdam: John Benjamins. 73–82.

Shreve, G., Lacruz I. and Angelone, E. 2011. Sight translation and speech disfluency. Performance analysis as a window to cognitive translation processes. In: Alvstad, C., Hild, A. and Tiselius, E. *Methods and Strategies of Process Research: Integrative approaches in Translation Studies.* Amsterdam: John Benjamins. 93-120.

Filled and unfilled pauses in recurrent multiword sequences in L2 presentational talk

Nicole BAUMGARTEN

University of Sheffield

The paper presents results from a longitudinal study of L2 development and register acquisition in English academic speaking in an internationalized university environment outside a native English speaking context, the so-called expanding circle of English language use (Kachru, 1985). In the environment investigated, English figures both as a means of lingua franca communication and as learning objective in dedicated English LSP courses. It has been suggested that in this kind of environment, language use operates under conditions of reduced common ground: There is limited access to a consistently norm-providing framework that regulates situational appropriateness of linguistic choice and participants experience heightened variability in language forms, form-function mappings and interactional choice (Le Page and Tabouret-Keller, 1985). It has been suggested further that under these circumstances normativity is suspended so that non-standard choice is not made interactionally relevant (Firth, 2009). The resultant unpredictability of linguistic choice opens up for L2 speakers a multitude of L2 learning trajectories and socialization paths. When communicative situations and constellations recur, the groups of speakers may evolve into a community of practice (Lave and Wenger, 1998), including the acquisition and development of a linguistic repertoire through potentially idiosyncratic appropriation of the L2 (Kramsch, 1998).

The present paper investigates recurrent multiword sequences (RMS) as markers of L2 development with a focus on the role and functions of filled and unfilled pauses in L2 speakers' RMS repertoire and use in order to gain insight into the development of routinized L2 discourse production in an expanding circle learning and usage environment. RMS have been investigated under a variety of names, including lexical phrases (Nattinger and DeCarrico, 1992), routine formulae (Coulmas, 1979), conversational routines (Aijmer, 1996), recurrent word combinations (Altenberg, 1998), lexical stems (Pawley and Syder, 1983), lexical bundles (Biber et al., 2004), formulas (Simpson-Vlach and Ellis, 2010), and formulaic sequences (Wray, 2002).

Their use has increasingly come to be considered as an important measure of L2 development in adults' spoken English and as a prerequisite for fluent L2 delivery (Adolphs and Durow, 2004; Crossley and Salisbury, 2011; De Cock, 2000; Wood, 2006; Ellis et al., 2008; Qi and Ding, 2011; Sánchez-Hernández, 2013). The quantitative presence of RMS in a speaker's speech reflects the degree of routinization of discourse production (Altenberg, 1998; Biber et al., 2004) as RMS are used as building blocks of and scaffolding devices for continuous discourse. Relatively little is known, however, about the development of fluency through acquisition of RMS in environments that are characterized by limited range and depth of target language use (Baumgarten 2014).

In addition, as far as can be gleaned from the descriptions of methodology in the research on L2 acquisition and use of RMS to date, the presence of filled and unfilled pauses within and around RMS has not been systematically considered in discussions of L2 speakers' capacity to produce uninterrupted 'chunks' or stretches of speech. As De Cock (2000) suggested, however, rather than diminishing over time, filled pauses might come to be analysed by L2 speakers as parts of a multiword sequence. For L1 collocation use, Kjellmer (2003) has already shown that *er* and *erm* can be functional parts of collocations. This is in line with research on filled (in particular hesitation makers *uh* and *um*) and unfilled pauses, which revealed that they function as markers of discourse structure and intersubjectivity, and carry non-linguistic, social meaning (e.g. Stenström, 1994; Swerts 1998; Mukherjee 2000; RendleShort, 2006; Tottie 2011).

The present study explores the L2 development of RMS and the functional status of filled and unfilled pauses within and around them through a longitudinal study of 10 students' L2 performance in the academic register *oral presentation*. We chose to focus on this type of speaking because academic presentations have been described as high stakes communicative events for L2 speakers: They require L2 production over a prolonged period of time under real-time production constraints (Graham & Barone, 2001; Nesi & Basturkmen, 2009). As such, academic presentations are instances in which pre-patterned language and routinization of discourse production help speakers to manage the affordances of the presentation situation. The presentations in our data occurred as regular monologic speaking assignments in English LSP classes. They were sampled on three naturally occurring classroom occasions during the students' first, second and final year of undergraduate study. The presentations were videorecorded and manually transcribed using an orthography-based transcription model (Rehbein et al., 2004). The RMS in students' talk were identified using the *clusters/ngrams*-function of the concordancer software AntConc (Anthony, 2012) and analysed for their structure, functional diversification, frequency, and changes to structure, function and frequency over time across individual L2 speakers and in the group. The analysis revealed that rather than just decreasing with time and exposure to the register, the pauses become constitutive parts of the RMS and contribute to their functions in discourse. These functions relate to discourse structuring, the expression of intersubjectivity, and speaker self-management in the presentation situation where RMS functions as stock turn holders to allow online planning and express phases of speaker deliberation at the same time. These results indicate that in monologic academic speaking pauses become part of what speakers themselves may experience as fluent, automatized and routinized L2 delivery and increasingly expert performance in the register. This, in turn, means that pedagogical interventions to eradicate disfluencies caused by pauses and hesitations will initially disrupt speakers' subjective experiences of fluent speech production and habitual performance patterns. Further research needs to clarify the register-specificity of RMS and pause patterns in both L1 and L2 use to determine whether L2 RMS

and pause patterns are a feature of learner language – and possibly the effect of a poverty of L2 exposure (Wray, 2000) in non-target language settings – or a function of the affordances of the speech situation that is unrelated to L1 or L2 speaker status.

Adolphs, Svenja and Valerie Durow. 2004. "Social-cultural integration and the development of formulaic sequences." *Formulaic sequences: Acquisition, processing and use*. Ed. Norbert Schmitt. Amsterdam: John Benjamins. 107-126.

Aijmer, Karin. 1996. *Conversational routines in English: Convention and creativity*. London/New York: Longman.

Altenberg, Bengt. 1998. "On the phraseology of spoken English: The evidence of recurrent word combinations." *Phraseology: Theory, analysis and applications*. Ed. A. P. Cowie. Oxford: Oxford University Press. 101-122.

Anthony, Laurence. 2012. AntConc. [AntConc 3.3.5w]. http://www.antlab.sci.waseda.ac.jp/software.html

Baumgarten, N. 2014. "Recurrent multiword sequences in L2 English spoken academic discourse: Developmental perspectives on 1st and 3rd year undergraduate presentational speech." *Nordic Journal of English Studies* 13/ 3: 1-32.

Biber, Douglas, Susan Conrad and Viviana Cortes. 2004. "*If you look at …*: Lexical bundles in university teaching and textbooks." *Applied Linguistics* 25/3: 371-405.

Coulmas, Florian. 1979. "On the sociolinguistic relevance of routine formulae." *Journal of Pragmatics* 3/3-4: 239-266.

Crossley, Scott Thomas Lee Salisbury. 2011. "The development of lexical bundle accuracy and production in English second language speakers." *International Review of Applied Linguistics* 49: 1-26.

De Cock, Sylvie. 2000. "Repetitive phrasal chunkiness and advanced EFL speech and writing". *Corpus Linguistics and Linguistic Theory. Papers from the Twentieth International Conference on English Language Research on Computerized Corpora (ICAME 20) Freiburg im Breisgau 1999*. Eds. Christian Mair and Marianne Hundt. Amsterdam: Rodopi. 51-68.

Swerts, Marc 1998. "Filled pauses as markers of discourse structure." Journal of Pragmatics 30: 485496.

Tottie, Gunnel (2011). "*Uh* and *Um* as sociolinguistic markers in British English" *International Journal of Corpus Linguistics* 16/2: 173–197.

Mukherjee, Joybrato 2000. "Speech is silver, but silence is golden: Some remarks on the function(s) of pauses". *Anglia*, 118/4: 571–584.

Ellis, Nick C., Rita Simpson-Vlach and Carson Maynard. 2008. "Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL." *TESOL Quarterly* 42/3: 375-396

Firth, Alan. 2009. "The lingua franca factor." *Intercultural Pragmatics* 6/2: 147-170.

Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: the English language in the outer circle. *English in the world: teaching and learning the language and literatures.* Eds. Randolph Quirk and Henry G. Widdowson. 11-30.

Kjellmer, G. 2003. "Hesitation. In defence of *er* and *erm*". *English Studies*, 84/2: 170–198.

Kramsch, Claire. 1998. *Language and culture.* Oxford: Oxford University Press.

Lave, Jean and Etienne Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Le Page, Robert B. and Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.

Nattinger, James F. and Jeanette DeCarrico. 1992. *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

Nesi, Hilary and Helen Basturkmen. 2009. "Lexical bundles and discourse signalling in academic lectures." *Lexical cohesion and corpus linguistics*. Eds. John Flowerdew and Michaela Mahlberg. Amsterdam: John Benjamins. 23-44.

Pawley, Andrew and Frances Hodgetts Syder. 1983. "Two puzzles for linguistic theory: Nativelike selection and nativelike fluency." *Language & Communication*. Eds. Jack C. Richards and Richard W. Schmidt. London/New York: Longman. 191-226.

Qi, Yan and Yanren Ding. 2011. "Use of formulaic sequences in monologues of Chinese EFL learners." *System* 39: 164-174. Cambridge: Cambridge University Press.

Rehbein, Jochen, Thomas Schmidt, Bernd Meyer, FranziskaWatzke and Annette Herkenrath. 2004. "Handbuch für das computergestützte Transkribieren nach HIAT". *Arbeiten zur Mehrsprachigkeit, Folge B* 56: 1-90.

Rendle-Short, Johanna. 2006. *The Academic Presentation: Situated Talk in Action*. Ashgate Publishing Ltd, Aldershot UK.

Sánchez-Hernández, Purificación. 2013. "Lexical bundles in three oral corpora of university students." *Nordic Journal of English Studies* 12/1, 187-209.

Simpson-Vlach, Rita and Nick C. Ellis. 2010. "An academic formulas list: New methods in phraseology research." *Applied Linguistics* 31/4: 487-512.

Stenström, Anna Brita. 1990. "Pauses in monologue and dialogue." *The London-Lund corpus of spoken English: description and research*.211-252.

Wood, David. 2006. "Uses and functions of formulaic sequences in second-language speech: An exploration of the foundations of fluency." *The Canadian Modern Language Review / La revue canadienne des langues vivantes* 63/1: 13-33.

Wray, Alison. 2000. "Formulaic sequences in second language teaching: Principle and practice." *Applied Linguistics* 21/4: 463-489.

Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Synthesized lengthening of function words - The fuzzy boundary between fluency and disfluency

Simon BETZ, Sina ZARRIEß, Petra WAGNER

Bielefeld University

I. INTRODUCTION & BACKGROUND

As [1]'s model of speech production suggests, speakers sense upcoming difficulties and can correct them before uttering. A reasonable strategy to bridge resulting gaps is to prolong the words in the articulatory buffer [2]. This often buys enough time to correct the issue, resulting in standalone disfluent lengthening, after which fluency is resumed [3]. In case of more severe difficulties, the lengthening may be followed by other disfluencies such as silent or filled pauses or repetitions. Similar hesitation strategies might be useful in automatic speech production, e.g. for spoken dialogue systems that interact with human users and typically face a variety of challenges in natural language understanding and generation.

Lengthening is an ambivalent phenomenon in speech that seems to be located at the fuzzy boundary between fluency and disfluency. It regularly occurs before phrase boundaries [4][5] and besides constitutes a common hesitation disfluency. Some disfluencies consist of lengthening [3] only, and some lengthenings appear so subtle that they pass unnoticed [6][7].

We assume that these characteristics of lengthening make it a key component in spoken dialogue systems that are capable of producing disfluencies, as they enable to buy a variable amount of time whilst being unobtrusive to the listener [6]. It is not yet known, however, how much synthetic lengthening is acceptable and how lengthening influences the user's interaction with the system. To address these issues, this study tests the effects of step-wise increases of synthesized lengthening on user ratings and interaction speed.

II. METHODS

We designed a perception test to evaluate sound quality of lengthening. This test is embedded in a simple game, in which users are asked by a synthetic voice to move around pentomino pieces on a computer screen (figure 1). The instructions follow a fixed order of [<pick up a piece> <conjunction phrase> <move it onto another piece>] (cf. sentences in example 1 with the conjunction phrase in boldface). After each stimulus, to proceed, participants have to click one of the four quality feedback buttons that constitute a 4- point MOS-scale.

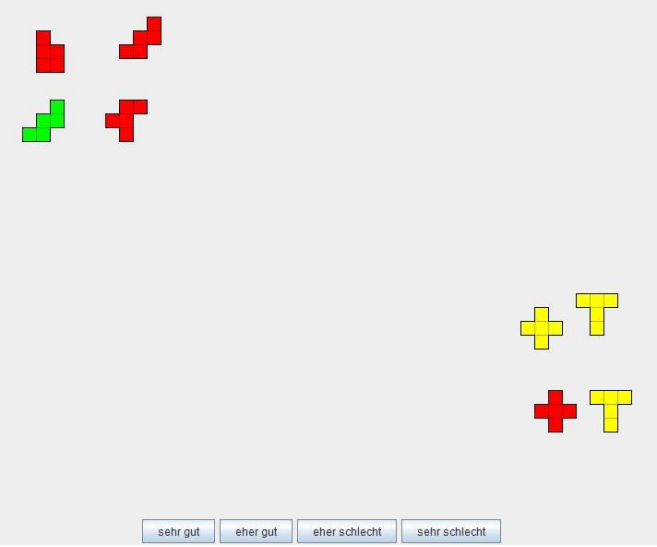


Fig. 1. Game scene with sound quality feedback buttons: very good, rather good, rather poor, very poor.

A. Stimulus design

Previous studies suggest that lengthening mainly occurs on function words [8][9], and that German articles, conjunctions and pronouns are frequent targets for lengthening [3]. For this study we test synthetic lengthening of function words in different degrees of lengthening with 400, 600, 800, 1000 , 1200 and 1400 ms duration of the target word. The target words are German monosyllables *(der, die, das, und, dann, ihn)* selected because of their high frequency of occurence and syllable-type balancing.[1] The duration for each segment in the target words is determined by applying the duration model based on the elasticity hypothesis [10], means and standard deviations for each phone are extracted from the GECO corpus [11].[1] Each target word

1 For this study, the "strong" form of the elasticity hypothesis was applied, i.e. general mean durations were used. At the moment, we test the reliability of an elasticity model that is based on disfluent lengthening data only to predict segment durations.

is embedded in a different carrier sentence and is located at the junction of two phrases that instruct the user to drag and drop pentomino pieces. The resulting six sentences (cf. Example 1) were synthesized in seven different configurations:

- The default configuration (i.e. with all segmental durations as predicted by the synthesizer's language model)
- The six different lengthening configurations (i.e. the same as the default, except that the target word's duration is set to 400, 600, ... 1400 ms.).

Example 1: Sentences (lengthened elements in boldface)

(1) *Nimm das rote Kreuz und lege es zum gelben Winkel.*

(2) *Die grune¨ Treppe, die muss ruber zum blauen Balken.*

(3) *Der gelbe Winkel, der muss ruber zum roten Balken.*

(4) *Das blaue Kreuz, das muss ruber zur grunen Treppe.*

(5) *Nimm die rote Treppe, dann lege sie zum gelben Kreuz.*

(6) *Nimm den grunen Balken und lege ihn zum blauen Winkel.*

In addition to the resulting 42 stimuli for analysis, we created 56 additional stimuli with different shapes and colors and without lengthening as distractors. Another six different stimuli were created for training the participants.

B. Stimulus presentation

Participants were instructed to act incrementally, i.e. start the task as soon as possible during the instruction and not wait until the voice has finished speaking. Each participant got the same set of 42 stimuli and 56 distractor sentences in a random order. Each session started with a short training phase to get participants used to the task.

C. Participants

23 participants took part in the experiment, all of them were students of Bielefeld University, between 19 and 37 years old (mean age 26.3). Six of the participants (26%) were male, 16 (73%) female and one of other gender. 20 (86%) had German as their mother tongue. 15 (66%) had previous experience with some kind of speech synthesis. None reported impairments of vision or hearing. The participants were paid 3e for their effort. None of the above mentioned variables (gender, mother tongue, experience with synthesis) had any apparent influence on the results. One participant was excluded from the final analysis, because inspection of their data revealed that they did not proceed incrementally.

III. RESULTS

Following suggestions by [12], we used R [13] with the lme4 package [14] to conduct a linear mixed effects analysis of the influence of lengthening extent on user ratings. As fixed effect, we had lengthening extent. As random effects we had intercepts for stimuli and participants, as well as by-stimulus and by-participant random slopes for the effect of lengthening extent, to control for ideosyncrasies of the participants and stimuli. Visual inspection of the residuals did not reveal any obvious deviations of homoscedasticity or normality.

We found that regardless of stimulus and participant, lengthening extent influences user ratings ($t(743) = -6.855$) , each increment lowering the average rating score by about 0.18 ± 0.027 (standard errors), on a scale where 4 corresponds to the best and 1 to the worst rating.

In addition to the ratings, we measured relative task completion times and checked for influences of lengthening extent. To control for the different sentence lengths, we calculated

the time span from beginning of audio until the drop of the pentomino piece divided by sentence duration. Using the same mixed models approach as above, we found that lengthening also significantly lowers relative task completion times ($t(743) = -4.296$), indicating that participants are not confused by the lengthening, but rather use the extra time to complete the task.

IV. DISCUSSION

As can be seen in Fig. 2, stimuli get good overall feedback and the ratings decline very slowly as lengthening increases, reaching a sustained trough at 1200ms. On the one hand, this leads to the assumption that even relatively long lengthening is a valid strategy for spoken dialogue systems. On the other hand, it suggests that lengthening should ideally be kept low to maintain highest-possible quality. Analyses of the interaction speed support this assumption, cf. Fig. 3. Users use the extra time granted by lengthening to solve the task - they get faster relative to sentence duration as lengthening increases, but appear to get distracted by extreme lengthening, when they appear to slow down again (although the slowdown is not significant).

Even lacking any evidence for lengthenings >1200*ms*, we take these as indicators for a turning point in synthesis quality around 1200ms: In terms of ratings, users do not differentiate anymore; in terms of task completion times, users need more time.

We furthermore suspect that lengthening is sometimes hard to notice due to its frequency of occurrence and its diversity of functions in everyday speech [6][7]. Summing up, our results raise the question as to the point at which lengthening characterizes a disfluency. In this experiment, we deliberately operationalized lengthening as a means to express hesitation, so it certainly counts as a disfluency from the production perspective. However, we still do not know the exact point (or the exact extent of lengthening) at which listeners start perceiving it as a disfluency. The slow and steady decline of our ratings suggests a fuzzy boundary rather than a clear threshold between “fluent” and “disfluent” lengthening.

V. CONCLUSION

We showed that synthesized lengthening gets good user feedback and does not negatively impact interaction speed.

Although this study reveals more of a fuzzy boundary than a clear threshold in lengthening acceptability, ratings and interaction times in the conditions over 1000ms suggest that there is an upper limit to synthetic lengthening. Possible follow-ups could examine the impact of greater lengthening extents to determine whether there is a turning point around 1200ms or whether this is merely an outlier. Lengthening in general appears well suited for disfluency synthesis. It is to be determined if longer hesitations should be covered by lengthening over multiple words or with combinations with other disfluencies such as silent and filled pauses.

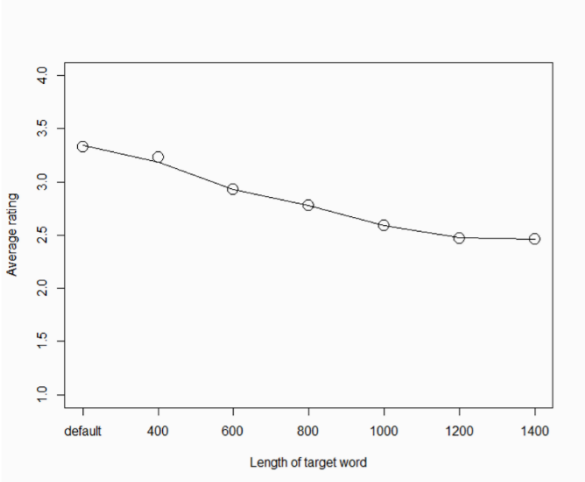


Fig. 2. User feedback with respect to word length. 4=good, 1=bad

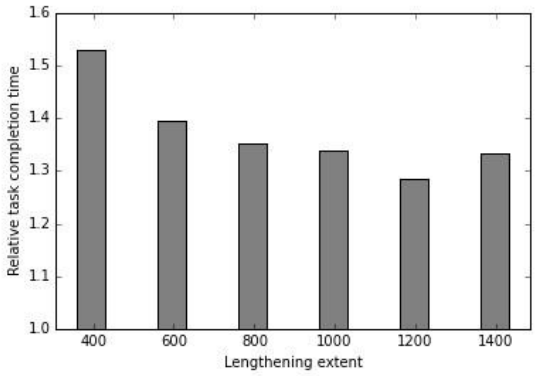


Fig. 3. Relative task completion time (divided by stimulus duration) over the different lengthening conditions

[1] W. J. Levelt, "Monitoring and self-repair in speech," *Cognition*, vol. 14, no. 1, pp. 41–104, 1983.

[2] J. Li and S. Tilsen, "Phonetic evidence for two types of disfluency," in *Proceedings of ICPhS 2015*, 2015.

[3] S. Betz, P. Wagner, and J. Vosse, "Deriving a strategy for synthesizing lengthening disfluencies based on spontaneous conversational speech data," in *Phonetik und Phonologie 12*, 2016.

[4] B. Peters, K. J. Kohler, and T. Wesener, "Phonetische Merkmale prosodischer Phrasierung in deutscher Spontansprache," *Prosodic structures in German spontaneous speech (AIPUK 35a)*, vol. 35a, pp. 143–184, 2005.

[5] A. E. Turk and S. Shattuck-Hufnagel, "Multiple targets of phrasefinal lengthening in American English words," *Journal of Phonetics*, vol. 35, no. 4, pp. 445–472, oct 2007.

[6] S. Betz, P. Wagner, and D. Schlangen, "Micro-structure of disfluencies: Basics for conversational speech synthesis," in *Proceedings of the 16th Annual Conference of the International Speech Communication Association (Interspeech 2015, Dresden)*, 2015, pp. 2222–2226.

[7] R. J. Lickley and E. G. Bard, "When can listeners detect disfluency in spontaneous speech?" *Language and speech*, vol. 41, no. 2, p. 203 , Apr 01 1998.

[8] E. Shriberg, "Preliminaries to a theory of speech disfluencies," *Ph D. thesis University of California*, 1994.

[9] D. O'Shaughnessy, "Timing patterns in fluent and disfluent spontaneous speech," in *International Conference on Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995*, vol. 1. IEEE, 1995, pp. 600–603.

[10] W. N. Campbell and S. D. Isard, "Segment durations in a syllable frame," *Journal of Phonetics*, vol. 19, no. 1, pp. 37–47, 1991.

[11] A. Schweitzer and N. Lewandowski, "Convergence of articulation rate in spontaneous speech," in *Proceedings of the 14th Annual Conference of the International Speech Communication Association (Interspeech 2013, Lyon), pp. 525–529.*

[12] B. Winter, Linear models and linear mixed effects models in R with linguistic applications., 2013. [Online].Available: http://arxiv.org/pdf/1308.5499.pdf

[13] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2015. [Online]. Available: https://www.R-project.org/

[14] D. Bates, M. Machler¨, B. Bolker, and S. Walker, "Fitting linear mixedeffects models using lme4," *Journal of Statistical Software*, vol. 67 , no. 1, pp. 1–48, 2015.

Disfluencies in children's and adolescents' spontaneous speech: the effect of speech task

Judit BÓNA, Tímea VAKULA

Eötvös Loránd University

Frequency of disfluencies in speech is affected by several factors like speech task and speaker's age. The main question of this presentation is if there are differences in the occurrence of disfluencies of children and adolescents depending on speech task.

80 subjects' recordings from GABI, the Hungarian Child Language and Speech Database and Information Repository (Bóna et al. 2014) were selected for the research: twenty 6-year-old, twenty 9-year-old children, and twenty 13-year-old and twenty 17-year-old adolescents. In all groups there were 10 males and 10 females. All of them were native Hungarian speakers with normal hearing, and didn't have any speech or language disorder.

Recordings were made with each subject in three situations which represented different speech tasks and required various cognitive skills with various levels of difficulty. 1. Spontaneous narrative (subjects spoke about their own lives); 2. narrative recalls (the task was to recall a short story they had listened to as accurately as possible; the success of it is determined by speech processing, attentional and working memory mechanisms, and narrative competence; Juncos-Rabadán–Pereiro 1999). 3. storytelling (the subjects had to relate a story on the basis of a series of six pictures.)

The frequency of every disfluency was defined for all speakers. Each occurrence and type of disfluencies were identified and coded by the two authors. The rate of agreement was 98% between the two coders. The types of disfluencies were analyzed, as follows (Searle et al., 2002; Roberts et al., 2009): interjections, revisions, repetitions (word- or phrase-repetitions and partword repetitions), and lengthenings. The data were compared across the age groups and the three speech tasks.

Preliminary results show that speakers' age has significant effect on the frequency of disfluencies, but there is significant difference between the speech tasks only in case of the 13-year-old and the 17-year-old groups.

Pedagogical implications: The results show which speech task is easier to produce for the children in the late stages of speech development and for the adolescents, and what language planning problems they have while speaking.

Bóna, J., Imre, A., Markó, A., Váradi, V., & Gósy M. (2014). GABI – Gyermeknyelvi beszédadatbázis és információtár. *Beszédkutatás 2014.* 246-251.

Juncos-Rabadán, O., and Pereiro, A. X. (1999). Telling stories in the elderly. Influence of attentional and working memory proceses (preliminary study), in *Psycholinguistics on the threshold of the year 2000. Proceedings of the 5th International Congress of the International Society of Applied Psycholinguistics*, edited by M. da Graça Pinto, J. Veloso, and B. Maia. (Faculdade de letras da Universidade do Porto), 155-159.

Roberts, P. M., Meltzer, A. & Wilding, J. (2009). Disfluencies in non-stuttering adults across sample lengths and topics. *Journal of Communication Disorders* 42. 414-427.

Searl, J. P., Gabel, R. M., & Fulks, J. S. (2002). Speech disfluency in centenarians. *Journal of Communication Disorders, 35.* 383-392.

Comparing the evaluation and processing of native and non-native disfluencies

Hans Rutger BOSKER[1,2]

Max Planck Institute for Psycholinguistics[1], *Radboud University*[2]

Introduction

Disfluency is a common feature of both native and nonnative speech. Nevertheless, there are considerable differences between native and non-native disfluency in the incidence and distribution of disfluencies. This study is concerned with the consequences of these differences between native and nonnative disfluency for *fluency perception*. If native and nonnative disfluencies are perceived differently by listeners, this could carry implications for the language testing practice since it may reveal whether the fluency evaluation of non-native speakers in language tests reflects the way we listen to native speakers in everyday communication.

Two research questions were formulated: (1) do listeners weigh native and non-native disfluencies differently in *fluency evaluation*?; and (2) do listeners process native and non-native disfluencies differently in *online speech processing*?

Experiments

The first two experiments concerned fluency evaluation. Semi-spontaneous speech excerpts of approximately 20 seconds from 10 native and 10 non-native speakers (intermediate proficiency) of Dutch were selected, with comparable numbers of silent pauses. In Experiment 1, the length of these silent pauses was manipulated resulting in three pause conditions: No Pauses (all silent pauses manipulated to have a duration <150 ms); Short Pauses (all silent pauses manipulated to have a duration between 250-500 ms); and Long Pauses (all silent pauses manipulated to have a duration between 750-1000 ms).

Participants in **Experiment 1** (*N*=58; naïve raters) listened to the manipulated native and non-native speech with instructions to rate the speech for fluency on a scale from 1 (very disfluent) to 9 (very fluent). Results revealed that longer pauses led to lower fluency ratings, as expected. Moreover, this effect was comparable for the native and the non-native speech. That is, lengthened pauses in native speech led to a similar decrease in perceived fluency level as lengthened pauses in non-native speech.

Participants in **Experiment 2** (*N*=57; new naïve raters) listened to the native and non-native speech in two rate conditions: Original (no rate manipulation); and Transplanted (native speech was played at a relatively slow non-native rate; non-native speech was played at a relatively fast native rate). Results showed that slowing down native speech led to a decrease in perceived fluency level, and speeding up nonnative speech led to an increase in perceived fluency level.

Moreover, the increase and decrease induced by the rate manipulations was of similar magnitude across native and non-native speech. Taken together, Experiment 1 and 2 suggest that native and non-native disfluencies are weighed in a similar fashion.

The final two experiments concerned the *online processing* of native and non-native disfluencies. **Experiment 3** involved an eye-tracking experiment which investigated the processing of filled pauses (*uh*) in native and non-native speech. Since filled pauses in native speech typically occur before low frequency words [1], we argued that listeners may anticipate reference to a low frequency object upon hearing a native disfluency. Conversely, in non-native speech filled pauses have a more irregular distribution [2], potentially making non-native *uh*'s worse predictors of the word to follow.

Participants in Group A (*N*=35) were presented with fluent and disfluent versions of referring expressions like *Click on thee uh…* [Klik op uh de …], recorded from a native speaker of Dutch,

followed by a high frequency target (e.g., *bike*) or a low frequency target (e.g., *sewing machine*). Participants in Group B (*N*=36) heard the same referring expressions, only this time recorded from a non-native speaker of Dutch (low proficiency; strong foreign accent). Participants' task was to click on one of two objects representing the sentence-final target presented visually on a computer screen, while their eye fixations were recorded using eye-tracking.

Results showed that, in fluent sentences, participants' did not have a preference for either object on screen. In disfluent speech, participants in Group A showed anticipatory looks towards low frequency objects, suggesting that they used the native disfluency to predict a low frequency referent. However, participants in Group B listening to non-native speech did not show this anticipatory bias. These findings demonstrate that native disfluencies elicit prediction of more complex referents whereas non-native disfluencies do not. This may potentially be explained by the difference in disfluency distributions in native and non-native speech, with non-native disfluencies being worse predictors of the word to follow, thus attenuating listeners' anticipatory behavior.

Finally, **Experiment 4** investigated the processing of articulation rate in native and non-native speech. The perceived rate of a precursor sentence may influence the perception of subsequent vowels. For instance, the perception of a Dutch vowel ambiguous between short /ɑ/ and long /a:/ may be biased towards /a:/ by presenting it after a fast precursor sentence [3]. To compare this rate effect across native and non-native speech, participants in Experiment 4 (*N*=45) were presented with fast and slow precursor sentences (recorded from 2 native and 2 non-native speakers), followed by target words ambiguous between /ɑ/ and /a:/. The acoustic characteristics of the native and non-native speakers were matched (identical fast and slow articulation rates, identical target vowels, etc.). Results showed that, indeed, fast precursors biased target perception towards /a:/. Moreover, in non-native speech there was an additional overall bias towards /a:/, suggesting that non-native speech was perceived as perceptually faster than matched native speech.

Discussion

Our findings show that native and non-native disfluencies are weighed similarly in fluency evaluation. This suggests that fluency evaluation of non-native speakers in language tests is a valid method reflecting everyday communication with native speakers. Nevertheless, the processing of non-native disfluencies is different from the processing of native disfluencies: non-native *uh*'s do not elicit prediction of low frequency referents and non-native articulation rate is perceived as faster than matched native speech, showing that 'explicit' fluency ratings and 'implicit' perception dissociate. As such, fluency evaluation is not informative of the online cognitive processes involved in speech comprehension, but only of their outcome.

[1] W. J. M. Levelt, "Monitoring and self-repair in speech," *Cognition*, vol. 14, pp. 41-104, 1983.

[2] A. Davies, *The native speaker: Myth and reality*. Clevedon: Multilingual Matters, 2003.

[3] H. R. Bosker, "Accounting for rate-dependent category boundary shifts in speech perception," *Attention, Perception & Psychophysics*, pp. 1-11, 2016.

You know in learner English as an aid to fluency and beyond

Lieven BUYSSE

KU Leuven

Pragmatic markers have been shown to function at various levels of the discourse (see e.g. Schiffrin, 1987; Redeker, 1990; Aijmer, 2002; Fischer, 2006) with two major types of functions, viz. structural (or textual) functions and interpersonal (or involvement) functions. Especially in their former capacity, pragmatic markers can be considered as providing "the 'grease' between parts of discourse" (Aijmer, 2013: 31). As such they contribute to the fluency of spoken discourse and may help speakers overcome disfluency issues.

Over the past few decades the surge of scholarly interest in pragmatic markers has also addressed non-native speaker perspectives. Such studies for English have brought to light differences between native speakers and learners. These largely consisted in reported "underuse" of a broad range of pragmatic markers by the learners (e.g. Romero Trillo, 2002; Hellermann and Vergun, 2006; Müller, 2005), with only a few exceptions (e.g. Müller, 2005, Aijmer, 2011 and Buysse, 2015 on *well*; Buysse, 2012 on *so*). One of the markers that does not appear to receive much attention from learners is *you know*. This pragmatic marker has often been described as primarily a marker of common ground, fostering relations between co-participants and serving as a cue for a hearer to retrieve an inferred message from assumed common ground (Östman, 1981; Jucker and Smith, 1998; Fox Tree and Schrock, 2002; Beeching, 2016). It has been argued that learners tend to shy away from pragmatic markers that fulfil such interpersonal functions (cf. Romero Trillo, 2002; Buysse, 2011). Another of *you know*'s most widely recognized functions is that of an editing marker (Holmes, 1986; Stubbe and Holmes, 1995; Erman, 2001; Fox Tree and Schrock, 2002; Müller, 2005): it occurs at moments where a speaker is looking for the right word or content, or needs to repair a prior utterance. In other words, when a speaker is experiencing problems processing or monitoring the discourse they may turn to *you know*. This would make *you know* precisely a suitable candidate to aid learners in signalling or overcoming such difficulties. In short, one major function of *you know* may have learners of English avoid this marker whereas another may on the contrary appeal to them and cater for their conversational needs.

The present study sets out to investigate how *you know* is used by learners of differing mother tongue backgrounds. Four components of the Louvain International Database of Spoken English Interlanguage (LINDSEI) will be examined. This corpus consists of interviews with learners of English approaching the final stages of their formal language learning process, as they are all majoring in English at university. These informal interviews were conducted by members of the teaching staff at the students' institutions according to a set format (Gilquin et al., 2010): the interviewee first talked for a few minutes about a topic (e.g. a film, book, travel experience or a life-changing experience), which sparked a conversation with the interviewer; at the end of each interview the learners were asked to tell a short story based on four pictures. For the purposes of this study the Dutch, French, German and Spanish components of LINDSEI were selected, each consisting of 50 interviews. The findings from this learner corpus are compared with those from the Louvain Corpus of Native English Conversation (LOCNEC), the native speaker reference corpus of LINDSEI that was compiled along the same lines as LINDSEI.

The qualitative analysis constitutes an in-depth corpus-driven functional analysis that demonstrates that all of the functions that appear in the native corpus also surface in the learner corpora. I will also contend that, regardless of whether it occurs as an editing marker or functions more on an overtly interpersonal level, *you know* always contributes to the fluency of

the speaker's discourse and of the conversation. Differences between the sub-corpora are particularly found in the quantitative analysis, where it is shown that (i) the incidence of *you know* is considerably lower in all learner sub-corpora than in the native sub-corpus, (ii) *you know* is considerably less frequent in the German sub-corpus than in the other learner sub-corpora, and (iii) within each sub-corpus there is much variation between individual speakers in the frequencies with which they turn to *you know*.

Aijmer, K. (2002) *English Discourse Particles. Evidence from a Corpus*. Amsterdam/Philadelphia: John Benjamins.

Aijmer, K. (2011) *Well I'm not sure I think…* The use of *well* by non-native speakers. *International Journal of Corpus Linguistics 16*(2), 231-254.

Aijmer, K. (2013) *Understanding Pragmatic Markers. A Variational Pragmatic Approach*. Edinburgh: Edinburgh University Press.

Beeching, K. (2016) *Pragmatic Markers in British English. Meaning in Social Interaction*. Cambridge: Cambridge University Press.

Buysse, L. (2011) The business of pragmatics. The case of discourse markers in the speech of students of Business English and English Linguistics. *ITL: International Journal of Applied Linguistics, 161*, 1030.

Buysse, L. (2012) *So* as a multifunctional discourse marker in native and learner speech. *Journal of Pragmatics, 44*(13), 1764-1782.

Buysse, L. (2015). 'Well it's not very ideal...' The pragmatic marker *well* in learner English. *Intercultural Pragmatics, 12*(1), 59-89.

Erman, B. (2001) Pragmatic markers revisited with a focus on *you know* in adult and adolescent talk. *Journal of Pragmatics 33*(9), 1337-1359.

Fischer, K. (ed.) (2006) *Approaches to Discourse Particles*. Amsterdam: Elsevier.

Fox Tree, J. E. and Schrock, J. C. (2002) Basic meanings of *you know* and *I mean*. *Journal of Pragmatics 34*(6), 727-747.

Gilquin, G., De Cock, S. and Granger, S. (2010) *Louvain International Database of Spoken English Interlanguage*. Louvain-la-Neuve: Presses Universitaires de Louvain.

Hellermann, J. and Vergun, A. (2007) Language which is not taught: The discourse marker use of beginning adult learners of English. *Journal of Pragmatics 39*(1), 157-179.

Holmes, J. (1986) Functions of *you know* in women's and men's speech. *Language in Society 15*(1), 121

Jucker, A. and Smith, S. (1998) And people just you know like 'wow': Discourse markers as negotiating strategies. In A. H. Jucker and Y. Ziv (eds.), *Discourse Markers: Descriptions and Theory*, pp. 171201. Amsterdam/Philadelphia: John Benjamins.

Müller, S. (2005) *Discourse Markers in Native and Non-Native English Discourse*. Amsterdam/Philadelphia: John Benjamins.

Östman, J.-O. (1981) You know: A discourse-functional approach. Amsterdam: John Benjamins.

Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics 14*(3), 367-381.

Romero Trillo, J. (2002) The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics 34*(6), 769-784.

Schiffrin, D. (1987) *Discourse Markers*. Cambridge: Cambridge University Press.

Stubbe, M. and Holmes, J. (1995) *You know, eh* and other 'exasperating expressions': An analysis of social and stylistic variation in the use of pragmatic devices in a sample of New Zealand English. *Language & Communication 15*(1), 63-88.

A preliminary contrastive analysis of the acoustic features of "nasal grunts" in the CID and NECTE corpora

Aurélie CHLÉBOWSKI

Université Paris Diderot, USPC

Introduction

This paper proposes a preliminary contrastive analysis of the meanings conveyed by "nasal grunts", "*i.e.* words which have no 'clear meaning' (Ward, 2000: 29) but possess a nasal feature" (Chlébowski and Ballier, 2015: 54).

We replicated on the CID corpus (Bertrand *et al*, 2008) – a corpus of spontaneous French – the study conducted by Chlébowski and Ballier (2015) and Chlébowski (2015) on the "nasal grunts" of Geordie English in the NECTE corpus (Allen *et al*, 2007). Their study included an experimental investigation of those sounds in an attempt to characterise an exhaustive typology of their acoustic features, *i.e.* "phonetic and prosodic components of 394 occurrences" (Chlébowski, 2016: 44), to which "they ascribed attitudinal meanings, according to the literature […] following a 'compositional model' (Ward, 2006: 55)" (Chlébowski, 2016: 44). They finally verified those meanings through a lexical-contextual analysis (Chlébowski and Ballier, 2015) and a preliminary perceptual evaluation (Chlébowski, 2016).

Following these authors, we hypothesise that "nasal grunts" in French convey interactional and attitudinal meanings encoded in their acoustic components. Therefore, beyond the fact that they may be categorised as *fillers, backchannels* or *disfluencies*, they would convey the same kind of meanings as words and - as regards interactions – express that the speaker is still fluent.

Method

The CID corpus (Bertrand *et al,* 2008) is composed of 8h of dialogue recordings involving 16 participants (10 female and 6 male natives of French language) performing two tasks of spontaneous language: 1) talk about professional conflicts and, 2) talk about odd situations (Bertrand *et al*, 2008: 3). Unlike the NECTE corpus (Allen *et al*, 2007) which was ecological (Chlébowski and Ballier, 2015: 54), the participants of the CID corpus (Bertrand *et al,* 2008) have been recorded using headsets microphones which reduce the background noises and enable us to analyse "nasal grunts" even when speakers talk at the same time (Bertrand *et al,* 2008: 4).

Investigating the "*phoneme* TextGrids" of this corpus with *AntConc software* (Anthony, 2012), we have found there were 2159 "nasal grunts" in the CID corpus (Bertrand *et* al, 2008), *i.e.* 424 of the type *hein,* 22 *han,* 166 *hum,* 1545 *mh,* and 2 *hm*. We have realised an experimental investigation with *Praat software* (Boersma and Weenink, 2009) of 4 files (*i.e.* AB, CM, EB and SR) from the CID corpus (Bertrand *et al,* 2008), *i.e.* 2 conversations of 2h each, from 2 male and 2 female speakers, for a total of 361 "nasal grunts" analysed out of 2159 (*i.e.* 56 *hein,* 4 *han,* 7 *hum,* 293 *mh* and 1 *hm*).

We have then extract automatic measurements of the duration (ms) and formants of those grunts by running a script with *Praat software* (Boersma and Weenink, 2009), analysed their register (Hz) semiautomatically and their fundamental frequency (ST), voice qualities (*i.e. creakiness* and *breathiness* or medial [h]), the presence of medial glottal stop and syllabification manually. This investigation showed that there were in fact 237 "nasal grunts" in the 4 files analysed, *i.e.* 21 grunts of the type /ɛˀ/, 7 /ɛː ɛˀ/, 3 /ɛ.h̃ ɛ̃ˀ/, 4 /inhaled aˀ/, 14 /inhaled m/, 46 /m/, 21 /m̩/, 16 /m.hm/, 68 /m.m/, 1 /m̩ .m̩/, 3 /m.m.m/, 1 /m.mh/, 14 /œm/, 4 /œm̩/, and 13 /œ̞ m̩/. This investigation also showed that just as the cases of clitics (Chlébowski and Ballier, 2015:

54) from the NECTE corpus (Allen *et al*, 2007), 39 of the grunts analysed here from the CID corpus (Bertrand *et al*, 2008) were linked to the previous words (*e.g. quand même_hein*, or *d'accord_euh_mh*) or the following words (*e.g. mh_ouais*). Moreover, 14 of the grunts analysed here were in fact part of laughter. All of "these cases were excluded, given our compositional perspective" (Chlébowski and Ballier, 2015: 54), but cases of the form *c'est fini_hein* were kept because they were deemed to be outside the phonological domain of clitic groups because of the hiatus, as opposed to cliticlike uses of grunts that may not add one syllable but may trigger a *liaison* (linking).

We have finally ascribed semantic meanings to each of the acoustic components found in this preliminary investigation of the CID corpus (Bertrand *et al*, 2008), based on French literature of French language when possible *e.g.* a high-rising contour would suggest that the speaker is asking a question (Morel and Danon-Boileau, 1998), and [œ] in *eum*, that the speaker is signalling to his/her interlocutor a problem is his/her speech and the wish to keep the floor (Candea, 2000). If not, we decided to keep those ascribed by Chlébowski and Ballier (2015) and Chlébowski (2015) on Geordie English (*e.g.* [m] would mean that the speaker is processing something while speaking (Chlébowski and Ballier, 2015: 55). We have then performed a lexical-contextual analysis of those meanings to confirm our intuitions.

Results

Results of the acoustic analysis showed that we can find the same acoustic components on French in the CID corpus (Bertrand *et al*, 2008) that Chlébowski and Ballier (2015) and Chlébowski (2015) found on Geordie English in the NECTE corpus (Allen *et al*, 2007), *i.e.* three groups of duration, low-rises, low-falls, fall-rises, rise-falls, a low-register, medial /h/ or breathiness, creakiness, nasal vowels, nasalised vowels and nasal-bilabial consonants. The analysis of formants on the vowels of "nasal grunts" from the CID corpus showed that the vowel in *euh_mh* would always be of the form [œ], the vowel in *hein* of the form [ɛ̃] , and the vowel in *han* of the form [ɑ̃] . It also tended to show that segmental acoustic components of "nasal grunts" were subjected to the phonological inventory of the targeted language, *e.g. ehm* [ɛm] in Geordie English *vs. eum* [œm] in French, but still the meaning is posited to be comparable.

Moreover, we found new acoustic components that we did not find in the NECTE corpus (Allen *et al*, 2007), *i.e.* inhaled [h], final [h] or *breathiness*, three syllabled grunts, and high-falls, and some components that we did found in the NECTE corpus were missing, *e.g.* high-rises. However, considering the fact that the investigation of the "nasal grunts" from the NECTE corpus (Allen *et al*, 2007) was performed on half of the PVC files (PVC, Milroy *et al*. 1997) and that of the "nasal grunts" from the CID corpus (Bertrand *et al*, 2008) on only 4 out of 16 files, finding new elements and lacking others is not enough to claim that those components exist in French and not in Geordie English and vice versa. To support this assumption, we will take the example of the interjection *hein* in French, which has its own entry in some dictionaries. According to the French *Centre National de ressources textuelles et lexicales* (Pierrel, 2005) *hein* is often used *"to ask for more information"* ("*pour demander à l'interlocuteur de completer, expliciter une information"*) and according to our compositional model, will therefore be uttered with a high-rising tone, which would then exist upon French "nasal grunts".

The fact that we found a final [h] or *breathiness* made us think that medial [h] or *breathiness* (*i.e.* [m.hm]) found by Chlébowski and Ballier (2015) and Chlébowski (2015) is in fact not medial, but the attack of the second syllable, and that or final [h] or *breathiness* will be a coda (*i.e.* [m.mh]). Moreover, as stated by Chlébowski and Ballier (2015: 55), the meaning ascribed to one-syllabled grunt is that the speaker takes a speaker's role, and that ascribed to two-syllabled grunt that the speaker takes a listening role. This would suggest that grunts uttered with fall-rising and rise-falling contours – grunt of 2 syllables with the complex contours broken

on each syllable – always mean that the speaker is taking a listening role. However, we found in the CID corpus (Bertrand *et al*, 2008) grunts of 3 syllables with those complex contours broken on the second and third syllables. We therefore hypothesise that grunts of 2 syllables uttered with complex contours mean that the speaker takes a speaker's role (*i.e.* just as a grunt of 1 syllable uttered with a simple contour) and that grunt of 3 syllables uttered with complex contours mean that the speaker takes a listening role (*i.e.* just as a grunt of 2 syllables with a simple contour).

Finally, though the lexical-contextual analysis of the 237 "nasal grunts" from the 4 files analysed from the CID corpus (Bertrand *et al*, 2008) is still in making, results suggest that each of the acoustic features of "nasal grunts" in French in the CID corpus (Bertrand *et al*, 2008) possess an invariant semantic value contrary to the idea that "the phonetic sequences involved in either onomatopoeia or sound symbolism are clearly not to be considered semantic constituents" (Cruse, 1986 :35). Moreover, when compared with the investigations conducted by Chlébowski and Ballier (2015) and Chlébowski (2016), our results tend to show that the semantic values conveyed by the acoustic components of "nasal grunts" are robust across French and Geordie English.

Aijmer, K. (2004). Pragmatic markers in spoken interlanguage. *Nordic Journal of English Studies*, *3*(1), 173-190.

Allen, W., Beal, J., Corrigan, K., Maguire, W., and H. Moisl. (2007). *A linguistic "time-capsule": the Newcastle Electronic Corpus of Tyneside English. Creating and Digitizing Language Corpora.* New

York/Basingstoke, Hampshire: Palgrave-Macmillan, 2. Available from: http://research.ncl.ac.uk/necte/index.htm [Accessed on: 18.01.2017]

Anthony, L. (2011). AntConc (Version 3.2. 2)[Computer Software]. *Tokyo, Japan: Waseda University*. Available from : http://www.laurenceanthony.net/ [Accessed on: 18.01.2017]

Bertrand, R., Blache, P., Espesser, R., Ferré, G., Meunier, C., Priego-Valverde, B., & Rauzy, S. (2008). Le CID-Corpus of Interactional Data-Annotation et exploitation multimodale de parole conversationnelle. *Traitement Automatique des Langues*, *49*(3), 1-30.

Boersma, P. and D. Weenink. (2009). *Praat*. Available from: http://www.fon.hum.uva.nl/praat/

[Accessed on: 18.01.2017]

Candea, M. (2000). Les euh et les allongements dits «d'hésitation»: deux phénomènes soumis à certaines contraintes en français oral non lu. *XXIIIèmes Journées d'étude sur la Parole*, 73-76. Chlébowski, A. (2015). "Nasal grunts" in the NECTE corpus, an experimental investigation. Unpublished MA thesis. Université Paris VII, Denis Diderot, France.

Chlébowski, A. (2016). The Meaning of "Nasal Grunts" in the Necte Corpus. A Preliminary Perceptual Investigation. *Research in Language*, *14*(1).

Chlébowski, A., & Ballier, N. (2015). "Nasal grunts" in the NECTE corpus–Meaningful interactional sounds. In *proceedings of EPIP4*, 54-58. Prague, Czech Republic May 21-23

Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press

Milroy, J., Milroy, L., & Docherty, G. (1997). *Phonological Variation and Change in Contemporary Spoken English*. Newcastle upon Tyne: Dept of speech, University of Newcastle upon Tyne. Retrieved from http://research.ncl.ac.uk/necte/index.htm [Accessed on: 18.01.2017]

Morel, M. A., & Danon-Boileau, L. (1998). *Grammaire de l'intonation l'exemple du français*. Editions OPHRYS.

Pierrel, J.M. (2005). Centre National de ressources textuelles et Lexicales. Available from http://www.cnrtl.fr/ [Accessed on: 18.01.2017]

Ward, N. G. (2006). "Non-lexical conversational sounds in American English", *Pragmatics & Cognition* 14(1),129-182. Available from http://digitalcommons.utep.edu/cs_techrep/307 [Accessed on: 18.01.2017]

Methodologies for modelling silent pause length. insights on individual and situational Variation from a large-scale corpus study

George CHRISTODOULIDES, Iulia GROSMAN, Liesbeth DEGAND, Anne-Catherine SIMON

Université catholique de Louvain

The temporal organisation of speech can be studied by segmenting the speech signal into measurable components, as a sequence of articulated intervals and pauses. Silent pauses fulfil multiple functions (Zellner, 1994), ranging from the most basic (e.g. breathing, or pre-occlusive pauses linked to articulation), to prosodic functions (e.g. as an important correlate of prosodic phrasing, cf. Krivokapic, 2007; Simon & Christodoulides, 2016), and discursive/rhetorical functions (e.g. to indicate saliency and focus, cf. Duez, 1982).

The statistical analysis of silent and filled pause length presents a number of methodological challenges, as the typical distribution of pause durations is positively skewed; therefore the use of methods that rest upon the hypothesis of normality is not appropriate (Oehmen, 2010).

An alternative method is to study the distribution of the *logarithm* of pause durations. For example, Kirsner & Hird (2005) find a bimodal distribution of the log-transformed silent pause lengths in their corpus and posit that the first component distribution (short pauses) corresponds to articulatory processes, while the second component (medium-length pauses) corresponds to cognitive processes, including discourse segmentation. However, the method of log-transformation must be applied with caution, after establishing that the original pause length distribution is indeed bimodal. Campione & Véronis (2002) analysed 5 hours of read and spontaneous speech in five languages, and report a trimodal distribution of pause length, categorizing them as brief (less than 200 ms), medium (200 to 1000 ms) and long (over 1000 ms). They only found long (>1s) pauses in spontaneous speech, and reported that pauses follow a log-normal distribution globally and for each category. Demol et al. (2007) analysed a 4-hour corpus of three different speaking styles, in six European languages, finding that the "logarithmic duration of the pauses can be well approximated by a bi-Gaussian distribution" both in slow and in fast speaking rates; similar pausing strategies were found for all languages (Dutch, English, French, Italian, Romanian and Spanish). Goldman et al. (2010) studied a 40-minute French spoken corpus with 4 speaking styles (reading, narration, broadcast news and university lectures): they report a multimodal distribution of log-transformed pause length and propose to model silent pauses as a mixture of log-normal distributions.

Furthermore, research in perception and psycholinguistics has shown that perceived pauses do not correspond to physical pauses. This is a manifestation of a more general property of the human sensory system: the perception threshold is higher than the actual physical stimulus (Zellner, 1994: 43), and is modulated by previously presented stimuli. We could therefore envisage modelling pausing behaviour using a measure that normalises the duration of each silent pause based on local speech rate, including the length of silent pauses in its immediate context: i.e. using *relative* length values, rather than absolute length values, or the logarithm of absolute length values.

We seek to address these methodological questions through a large-scale corpus study and statistical analysis of the properties of different measures of silent pause length. We have compiled five phonetically aligned corpora of French speech: the LOCAS-F corpus (Martin, Degand, and Simon 2014), the C-Humour corpus (Grosman 2016), the Driving Simulator Cognitive Load corpus (Christodoulides 2016), the C-Phonogenre corpus (Prsir, Goldman, and Auchlin 2014), and the Rhapsodie corpus (Lacheret et al. 2014). The compilation covers 31 speaking styles, includes a total of 276 different samples, and its total duration is 17.4 hours

(186.895 tokens). As the corpus contains both monologues and dialogues, we are only focusing on pauses *inside* a speaker's turn (betweenspeaker gaps have been excluded from the analysis). The corpus compilation contains approximately 23.000 turn-internal silent pauses.

Figure 1 shows the distribution of four different measures of pause duration: absolute length; the base10 logarithm of absolute length; the relative duration of each silent pause; and the base-10 logarithm of the relative duration of each silent pause. Relative duration is defined as the length of a pause divided by the arithmetic mean of the length of neighbouring segments within a window of ±5 segments (including both syllables and pauses).

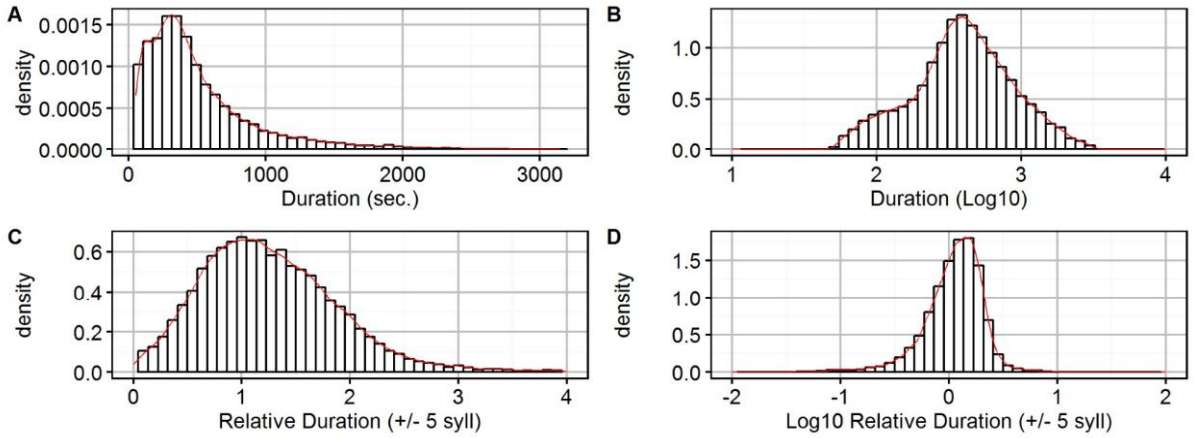


Figure 1 Distributions of four measures of silent pause length

As expected, none of the four measures follows a normal distribution. While panel B (log-transformed durations) may suggest that this measure produces a bimodal distribution (and described as a mixture of Gaussian models), we have applied these methods to each speaking style separately, and found that some speaking styles follow a unimodal distribution, others follow a bimodal distribution, and a few follow a trimodal distribution. Figure 2 shows the distribution of the log-transformed durations for four different speaking styles (academic speech, radio news, reading and stand-up comedy). We observe that, while $\log_{10}$(duration) appears to follow a bimodal distribution at the speaking style level (i.e. jointly modelling the distribution of pauses from several different speakers), this bimodality disappears at the individual speaker level. As can be seen in Figure 3, individual variation is greater in some speaking styles. Furthermore, the $\log_{10}$(relative duration) measure follows a unimodal distribution in almost all cases. These findings lead us to question the appropriateness of modelling silent pause length as a mixture of log-normal distributions. They also suggest that the local speech rate context may be playing a more important role than previously assumed (a hypothesis that should be tested through targeted perception experiments). Additional analyses will be presented, regarding the relationship between the statistical distribution of silent pause length, and (a) their position in the syntactical structure, as well as (b) their occurrence as part of a disfluency *interregnum*.

In this presentation we focus on methodological questions, on the basis of a corpus study. Perspectives for future research include applying these methods to describe the distribution of filled pause length. This corpus study will facilitate forthcoming perceptual experiments in order to validate which of the modelling methods is closer to the perception of silent pause length.

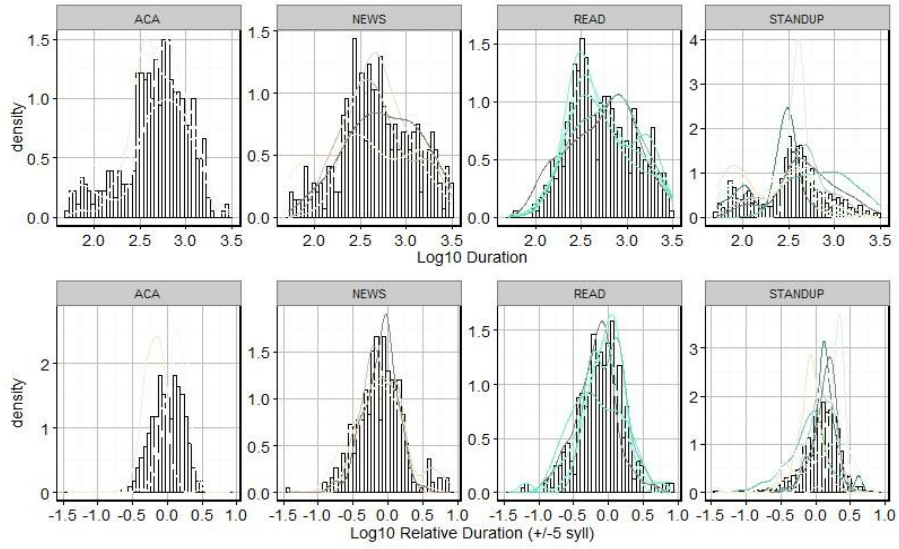


Figure 2 Differences in pause length distribution) by speaking style with individual variations

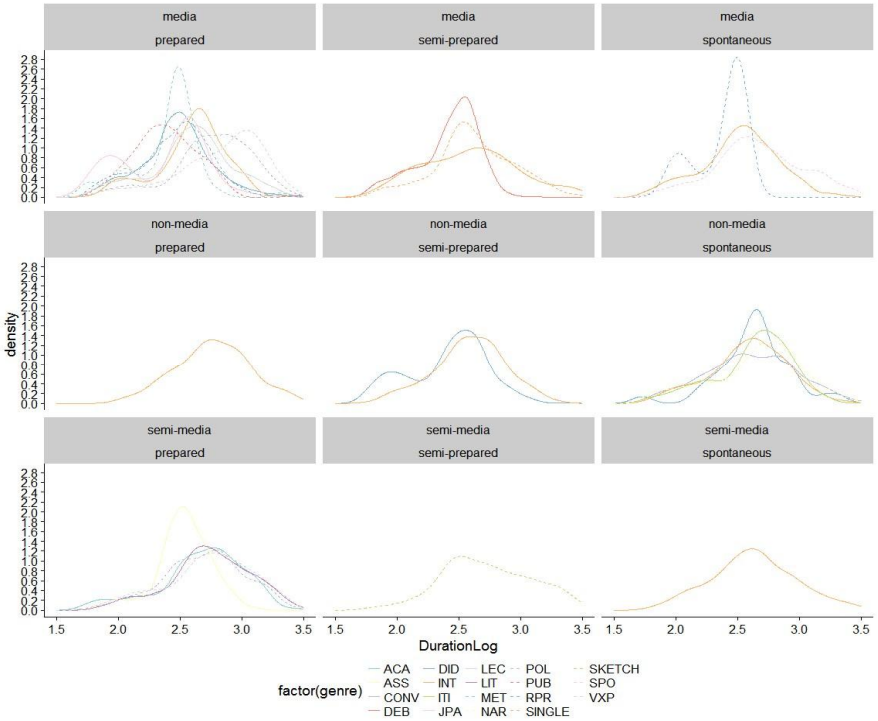


Figure 3 Pause length distributions by speaking style (grouped by two dimensions)

Campione, E. & Véronis, J. (2002). A Large-Scale Multilingual Study of Silent Pause Duration. In B. Bel & I. Marlien (Eds.), *Proceedings of Speech Prosody 2002* (pp. 199–202).

Christodoulides, G. (2016). Effects of Cognitive Load on Speech Production and Perception. PhD Thesis, Université catholique de Louvain.

Demol, M., Verhelst, W., & Verhoeve, P. (2007). The Duration of Speech Pauses in a Multilingual Environment. In *Proceedings of Interspeech 2007* (pp. 990–993).

Duez, D. (1982) Silent and Non-Silent Pauses in Three Speech Styles, *Language and Speech*, 25(1), 11–28.

Goldman, J.-Ph., François, T., Roekhaut, S., Simon, A. C. (2010). Étude statistique de la durée pausale dans différents styles de parole. XXVIIIèmes journées d'étude sur la parole (JEP 2010). In: *Actes des 28èmes journées d'étude sur la parole (JEP)*, Mons, 25-28 May 2010, 161-164.

Grosman, I. (2016). How Do French Humorists Manage Their Persona across Situations? A Corpus Study on Their Prosodic Variation. In: Ruiz-Gurillo, L. (Ed.) Metapragmatics of Humor: Current Research Trends (pp. 149–77). JB IVITRA. John Benjamins Publishing Company.

Kirsner, K., Dunn, J., & Hird, K. (2005). Language Production: a complex dynamic system with a chronometric footprint. In *Proceedings of the 7th International Conference on Cognitive Systems*.

Krivokapic, J. 2007. The planning, production, and perception of prosodic structure. PhD thesis, University of Southern California.

Lacheret, A., Kahane, S., Beliao, J., Dister, A., Gerdes, K., Goldman, J.-Ph., Obin, N., Pietrandrea, P., & Tchobanov, A. (2014). 'Rhapsodie: A Prosodic-Syntactic Treebank for Spoken French'. In *Proceedings of the 2014 Language Resources and Evaluation Conference (LREC)*. http://hal.upmc.fr/hal-00968959/.

Martin, L. J., Degand, L., & Simon, A.C. (2014). Forme et fonction de la périphérie gauche dans un corpus oral multigenres annoté. *Corpus*, Eléments initiaux dans la phrase : approches intergenres et inter-langues, no. 13: 243–65.

Oehmen, R., Kirsner, K., & Fay, N. (2010). Reliability of the manual segmentation of pauses in natural speech. In H. Loftsson, E. Rögnvaldsson, & S. Helgadóttir (Eds.), *Advances in Natural Language Processing* (Vol. 6233). Lecture Notes in Computer Science. Berlin, Heidelberg: Springer.

Prsir, T., Goldman, J.-Ph., & Auchlin, A. (2014). Prosodic Features of Situational Variation across Nine Speaking Styles in French. *Journal of Speech Sciences* 4 (1): 41–60.

Simon, A. C. & Christodoulides, G. (2016). Perception of Prosodic Boundaries by Naïve Listeners in French. In *Proceedings of Speech Prosody 2016*.

Zellner, B. (1994). Pauses and the temporal structure of speech. In E. Keller (Ed.), Fundamentals of speech synthesis and speech recognition (pp. 41–62). Chichester: Wiley.

Recycling discourse: from qualitative repair categories to a formal scale of fluency

Ludivine CRIBLE

Université catholique de Louvain

Spoken language is characterized by online processes of production and comprehension happening over time (Holšánová 2008). A natural consequence of this temporal nature is the presence of so-called disfluencies or "fluencemes" (Götz 2013) that can generally be defined as signals of on-going mechanisms of processing and monitoring. Authors (e.g. Shriberg 1994) usually include pauses, repetitions, discourse markers, truncations, substitutions and false starts. Fluencemes are omnipresent especially in impromptu speech, although not excluded from more prepared and monologic interactions. They relate to the complex concepts of fluency and disfluency, respectively marked by smooth vs. hesitant stretches of talk (e.g. Ejzenberg 2000). However, most disfluencies are ambivalent in their structure and function: a repetition of adjacent words can either signal lexical access trouble ("it's in the it's in the fridge") or stylistic emphasis ("it's very clear it's very clear"). This functional ambivalence as fluent or disfluent is, in principle, true for all fluencemes and connects the present approach with previous research showing the productive and strategic roles of (dis)fluent devices (e.g. Arnold et al. 2003, Corley et al. 2007): fluencemes are not only "symptoms" of hesitation but also "signals" for the listener (Clark & Fox Tree 2002).

This paper aims at bridging the gap between formal annotation (*how*) and pragmatic interpretation (*why*) of (dis)fluent devices in order to distinguish fluent from disfluent uses of the same structures and to propose a scale of (dis)fluency against which clusters of fluencemes can be evaluated. Concretely, stretches of adjacent fluencemes have been classified i) formally according to the fluencemes they contain and the relation between their different parts (e.g. immediate replacement of one word by another) and ii) functionally through a qualitative identification of the cause or motivation behind the repair (e.g. the speaker corrects an error in the first word). The resulting scale of (dis)fluency is based on this mapping of form and function, thus sheding new light on the role of specific fluencemes in the production of (dis)fluent discourse.

This study is strongly rooted in Levelt's (1983) model of speech production and monitoring. Levelt (1983) makes a basic distinction between error-correction and appropriateness-adjustment, and further identifies four types and six subtypes of repair: delay; error (lexical, syntactic, phonetic); appropriateness (terminological specificity, contextual ambiguity, lexical consistency); other. Levelt (1983) also includes other variables describing the format of the repair and the relation between the different parts of the repair such as moment of interruption or way of restarting. He successfully shows that, in Dutch, different repair types are expressed by different forms, in meaningful clusters of cues which are designed to help the listener interpret the utterance.

Levelt's (1983) categories have been revised and applied to 367 fluenceme clusters extracted from a French-English corpus of face-to-face interviews, compiled from the Valibel (Dister et al. 2009) and Backbone (Kohn 2012) corpora, respectively, where fluencemes had previously been identified and annotated following a multilingual and multimodal typology (Crible et al. 2016). Within this typology, the present study focuses in particular on the fluenceme of modified repetitions since it is particularly flexible and can be involved in all repair types from Levelt's (1983) model, as opposed to less ambivalent fluencemes such as identical repetitions or truncations, for instance. Examples 1–3 illustrate some of the uses of modified repetitions.

(1) a lot of them actually head down there head down to the Barbican and walk (EN-intf02)

(2) ils parlent mal le français euh (0.720) ou ils ont été mal éduqués en français (FRintf-01)
they don't speak French well uh (0.720) or they had a poor education in French

(3) the mums remember you and the dads remember you (EN-intf-03)

In (1), the words "head down" are repeated to introduce the change from the contextually ambiguous pronoun "there" to the specific referent "the Barbican" (appropriateness-repair, subtype: ambiguity). In (2), the speaker substitutes one verbal phrase with another (error-repair, subtype: lexical). In (3) however, we can see that all words are repeated except one ("mums" and "dads") but, unlike the previous two examples, this substitution is not motivated by an error or inappropriateness but rather creates an effect of resonance or parallelism, a category which was added to Levelt's (1983) original model.

Annotation of fluencemes and coding of repair types and format were mapped in order to find recurrent patterns in each language, following the hypotheses from Levelt (1983) and Fox et al. (1996) that i) detection of repair should occur sooner when the source of the repair is an error than when it is an issue of appropriateness; ii) within-word interruptions should only target erroneous words and not "neutral" words. Modified repetitions were also expected not to frequently co-occur with discourse markers (pragmatic expressions such as "actually", "ou",

"and" in the examples below) since their signaling function would be redundant and competitive with each other. Crosslinguistically, results from previous literature suggest that Romance languages make use of more complex and more ambiguous markers than English (Cuenca 2003), which should be reflected in more different types of discourse markers in French than in English repairs. Moreover, Auer & Pfänder (2007) found that French has a tendency to build parallel constructions with a rhetorical function, thus potentially resulting in more cases of "resonance" repairs as in Example (3).

The results show that lengthy repetitions seem to contribute to fluent hearer-oriented strategies, along with other features of repairs such as their syntactic integration in the original utterance or the structural or semantic resonances between segments, cues which are exclusively used in repair types of intermediate or high fluency. Another major finding is the attraction between modified repetitions and fluent repair types (i.e. the "resonance" category), while discourse markers appear rather absent from repairs in general and fluent repairs in particular, confirming my hypothesis on their redundancy – and therefore repulsion effect – with modified repetitions. Overall, the attempt to build a formal scale of (dis)fluency where different degrees of fluency are associated with objective features was successfully met by a number of cross-tabulations which converge in identifying the following three major patterns: repairs attending to utterance structure typically interrupt short units and introduce start-overs with fresh material (low fluency); repairs attending to lexical correctness are more integrated in the original utterance and cannot be formally divided into error-correction and appropriateness-adjustment (intermediate fluency); "resonance" repairs are strongly related to larger segments, long distances and high integration in the utterance (high fluency).

Arnold, J., Fagnano, M., & Tanenhaus, M. (2003). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research*, 32(1), 25-36.

Auer, P. & Pfänder, S. (2007). Multiple retractions in spoken French and spoken German. A contrastive study in oral performance styles. *Cahiers de praxématique* 48: 57-84.

Clark, H. & Fox Tree, J. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, 84, 73111.

Corley, M., MacGregor, L., & Donaldson, D. (2007). It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition*, 105, 658-668.

Crible, L., Dumont, A., Grosman, I. & Notarrigo, I. 2016. Annotation manual of fluency and disfluency markers in multilingual, multimodal, native and learner corpora. Version 2.0. *Technica report*. Université catholique de Louvain and Université de Namur.

Cuenca, M.J. 2003. Two ways to reformulate: a contrastive analysis of reformulation markers. *Journal of Pragmatics* 35: 1069-1093.

Dister, A., Francard, M., Hambye, P., & Simon, A.-C. (2009). Du corpus à la banque de données. Du son, des textes et des métadonnées. L'évolution de la banque de données textuelles orales VALIBEL (1989-2009). *Cahiers de Linguistique*, 33(2), 113-129.

Ejzenberg, R. (2000). The juggling act of oral fluency: a psycho-sociolinguistic metaphor. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 288-313). Ann Arbor, MI: The University of Michigan Press.

Fox B., Hayashi M., Jasperson R. (1996). Resources and repair : A cross-linguistic study of syntax and repair. In E. Ochs, E. Schegloff, S. Thompson (eds), *Interaction and grammar*. *Studies in interactional sociolinguistics* 13. 185-237.

Götz, S. (2013). *Fluency in native and nonnative English speech*. Amsterdam, Netherlands: John Benjamins.

Holšánová, J. (2008). *Discourse, Vision, and Cognition.* Amsterdam, Netherlands: John Benjamins.

Kohn, K. (2012). Pedagogic corpora for content and language integrated learning. Insights from the BACKBONE project. *The Eurocall Review*, 20(2).

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41-104.

Shriberg, E. (1994). *Preliminaries to a theory of speech disfluencies*. (Unpublished doctoral dissertation). University of California at Berkeley, CA.

(Dis)fluency across spoken and signed languages: applications of an interoperable annotation scheme

Ludivine CRIBLE[1], Amandine DUMONT[1], Iulia GROSMAN[1], Ingrid NOTARRIGO[2]

Université catholique de Louvain[1], Université de Namur[2]

Natural languages, whether spoken or signed, are characterized by the presence of so-called disfluencies or "fluencemes" (Götz 2013) which reflect the online nature of production and comprehension processes. Fluencemes are related to the complex constructs of fluency and disfluency, which are commonly associated with continuity and interruption, as illustrated in

Ejzenberg's (2000: 287) definition: "the degree to which speech is articulated smoothly and continuously without any 'unnatural' breakdowns in flow". Fluency is also fundamentally contextbound: the production and perception of fluencemes strongly depend on distributional (frequency, position, combination) and contextual factors (interactional features), so much so that the same fluenceme can either be fluent or disfluent depending on its local and situational environment. This ambivalence, acknowledged in our use of the term *(dis)fluency*, motivates a componential approach to the phenomenon that decomposes a holistic impression of "speed and effortlessness" (Chambers 1997: 535) into a typology of fluencemes which, once combined, allow the researcher to evaluate the relative fluency of the speaker or signer. Starting from Shriberg's (1994) seminal work, an increasing body of corpus studies have indeed adopted this componential view, though with slightly different definitions, methods and research agendas, across a variety of languages (e.g. English in Besser & Alexandersson 2007, Portuguese in Moniz 2013 or French in Pallaud et al. 2013).

Following this line of research, this paper presents an annotation scheme that aims to offer an exhaustive yet flexible coverage of fluencemes, and that is applicable to both spoken and signed languages. It was designed and tested on a variety of speech situations uttered in spoken French, native and learner English as well as in Belgian French Sign Language (henceforth LSFB). The fluencemes included in the model can be divided in four groups: (1) simple fluencemes – they include pauses (filled, unfilled, hand stops), palm-ups, discourse markers, explicit editing terms, false-starts and incomplete truncations; (2) compound fluencemes, a category that covers repetitions (identical, modified, framing) and substitutions (morphosyntactic, propositional); (3) insertions (lexical, parenthetical) and deletions; and (4) "diacritics" (the term is taken from Shriberg 1994), that add contextual information to an existing fluenceme (misarticulation, lengthening, embedding, re-ordering or syntactic completion). The main innovative aspects of this annotation scheme involve, first, its applicability to multilingual and multimodal corpora (especially to signed languages), and, second, its technical format that enables the handling of complex embedded structures on a

single layer of annotation thanks to a bracketing and numbering system. Another major, more theoretical, characteristic of the scheme concerns the ambivalent (fluent vs disfluent) nature of fluencemes: not only are the consensual hesitations and interruptions included in the typology, but typically "fluent" devices are also included, that is, the strategic uses of fluencemes for rhetorical or structuring effects, as in examples 1 and 2 below.

1. we have to have (0.980) um life saving floats we have to have (0.280) life buoys we have to have (0.470) bilge pumps (*Backbone* corpus, "bb_en002_amphibioustours")

2. AVOIR PT:DET DIRE BIEN AVOIR PT:DET ENVIRON (LSFB corpus, CLSFBI3705) *there are competent [interpreters] there are weaker [interpreters]*

These examples contain several occurrences of pauses and hand stops, as well as "fluent" modified repetitions ("we have to have" in English, AVOIR PT:DET in LSFB) which are not corrective reformulations but rather create an enumeration or a contrast, respectively, thus exploiting the linguistic material in the local context for strategic purposes. This, we argue, allows the researcher to refrain from early or arbitrary judgments at this stage of the analysis.

The replicability of the annotation was assessed by an inter-annotator agreement analysis on a 7.000word sample of radio interviews in spoken French, where we found substantial to almost perfect kappa-scores between two annotators: κ= 0.67 including disagreement on detection of fluencemes; κ = 0.82 when excluding disagreement on detection and categories with less than 10 occurrences. These results are very encouraging and indicate that the protocol can reliably be used in multilingual and multimodal corpora.

The potential of this annotation model is then illustrated in the genre of interviews in six corpora, namely Backbone (Kohn 2012), the French Corpus of Humorist Speech (C-Humour, Grosman 2016), the Louvain Corpus of Annotated Speech - French (LOCAS-F, Degand et al. 2014), the Louvain International Database of Spoken English Interlanguage (LINDSEI, Gilquin et al. 2010), the Corpus LSFB (Meurant 2015), and VALIBEL Database (Dister et al. 2009). The most frequent fluencemes were compared across the different languages and modalities, revealing whether signers and speakers of different languages and modalities resort to the same devices in a similar interaction. While spoken languages (i.e. native French, and native and learner English) show very similar quantitative behaviors (absolute and relative frequency of fluencemes), LSFB appears to make use of a different top set of fluencemes. For instance, modified repetitions are prevalent in LSFB and far more frequent than in spoken languages which prefer unfilled pauses: this result might be an indication of different strategies in coping with language production. In a last step, the study zooms in on the clustering tendencies of palm-ups and filled pauses (such as *uh* and *uhm* in English), investigating whether and under what conditions these two categories can be considered as multimodal equivalents, as is sometimes claimed in the literature (e.g. Locker McKee & Wallingford 2011; van Loon 2012). More specifically, palm-ups and filled pauses are systematically compared in terms of their proportion in isolation vs. in combination with other adjacent fluencemes and, if so, what and where these adjacent fluencemes appear (i.e. in initial, medial or final position in the sequence).

The method and analysis reported in this paper contribute to the growing interest for multimodal data, providing further insights into, on the one hand, the potential equivalence of fluencemes in the spoken and signed modalities and, on the other, how similar speaking situations might breed different outputs depending on (the nature of) the spoken language

(French/English, native/learner). Lastly, this multimodal and multilingual annotation scheme successfully addresses a number of methodological issues, namely universality and interoperability, in view of further enhancing our understanding of the complex notion of (dis)fluency.

Besser, J., & Alexandersson, J. (2007). A comprehensive disfluency model for multi-party interaction. In *Proceedings of the 8th SIGDial Workshop on Discourse and Dialogue*, Antwerp, Belgium.

Chambers, F. (1997). What do we mean by fluency? *System* 25(4), 535-544.

Degand, L., Martin, L., & Simon, A.-C. (2014). LOCAS-F: un corpus oral multigenres annoté. Paper presented at the *Congrès Mondial de Linguistique française*, Berlin, Germany.

Dister, A., Francard, M., Hambye, P., & Simon, A.-C. (2009). Du corpus à la banque de données. Du son, des textes et des métadonnées. L'évolution de la banque de données textuelles orales VALIBEL (1989-2009). *Cahiers de Linguistique*, 33(2), 113-129.

Ejzenberg, R. (2000). The juggling act of oral fluency: a psycho-sociolinguistic metaphor. In H. Riggenbach (Ed.), *Perspectives on fluency*, Ann Arbor, MI, The University of Michigan Press: 288313.

Gilquin, G., S. De Cock & S. Granger (eds.). (2010). *LINDSEI. Louvain International Database of Spoken English Interlanguage*. Presses Universitaires de Louvain. Louvain-la-Neuve.

Götz, S. (2013). *Fluency in native and nonnative English speech*. Amsterdam: John Benjamins.

Grosman, I. (2016). How do French humorists manage their persona across situations? A corpus study on their prosodic variation. In L. Ruiz-Gurillo (Ed.), *Metapragmatics of Humor: Current Research Trends*, Amsterdam, John Benjamins: 149-178.

Kurt, K. (2012). Pedagogic corpora for content and language integrated learning. Insights from the BACKBONE project. *The Eurocall Review*, 20(2).

Locker McKee, R. & Wallingford, S. (2011). "So, well, whatever" Discourse functions of palm-up in New Zealand Sign Language. *Sign Language and Linguistics* 14(2): 213-247.

Meurant, L. (2015). Corpus LSFB. First digital open access corpus of movies and annotations of

French Belgian Sign Language (LSFB). Laboratoire de langue des signes de Belgique francophone (LSFB-Lab). University of Namur. URL: http://www.corpus-lsfb.be.

Moniz, H. (2013). *Processing disfluencies in European Portuguese*. Doctoral dissertation, Universidade de Lisboa.

Pallaud, B., Rauzy, S., & Blâche, P. (2013). *Guide d'annotation. Identification et annotation des auto-interruptions et des disfluences dans les corpus du CID*. Unpublished technical report, Laboratoire Parole et Langage.

Shriberg, E. (1994). *Preliminaries to a theory of speech disfluencies*. Doctoral dissertation, University of California at Berkeley, CA.

van Loon, E. (2012). *What's in the palm of your hands? Discourse functions of Palm-up in Sign Language of the Netherlands*. Master thesis, University of Amsterdam.

Examining fluency in second language speaking from speaker's perspective: a cognitive approach

Tannistha DASGUPTA

The English and Foreign Languages University

This paper studies fluency in second language (L2) speaking from speaker's point of view i.e. the underlying mechanism involved in language production. Hence, it explores the aspects of L2 spoken fluency that are related to L2 linguistic knowledge and processing skill. The theoretical model adopted in the study is De Bot's (1992) adaptation of Levelt's (1989) 'blueprint' of monolingual speaker. The blueprint discusses the fundamental aspects like linguistic, psycholinguistic, and cognitive issues that underlie the act of speaking. In L2 speaking assessments, on the other hand, these aspects i.e. the underlying mechanism involved in language production, are generally disregarded (Van Moere, 2012). Usually, the relationship between subjective ratings on L2 speech and objectively measured aspects of fluency is studied (Lennon, 1990; Riggenbach, 1991; Freed; 1995; Kormos and D'enes, 2004; Rossiter, 2009). However, this study examines the relation between the underlying mechanism that is involved in L2 fluency in speaking, and L2 utterance fluency. In other words, it aims to find out the relation between L2 cognitive fluency (CF) and L2 utterance fluency (UF), where L2 cognitive fluency is defined as the efficiency of the underlying processes responsible for the producing utterances, and L2 utterance fluency is identified as the features of utterances that reflect L2 cognitive fluency (Segalowitz, 2010). In this study, the attempt is to find out which aspects of L2 utterance fluency are reliable indicators of L2 cognitive fluency.

The subjects of the study are 30 adult L2 learners of English from Hyderabad, India; with intermediate to lower-intermediate levels of proficiency in English. Data of the study includes scores in L2 linguistic knowledge test, results of lexical retrieval tasks and attentionshifting tasks (to measure CF), and 8 oral tasks (to measure UF). The three aspects of UF (speed, pause, & repair) are measured against the scores of CF. Bivariate Pearson correlation was employed to find out the predictive power of each aspect of UF (speed, breakdown, and repair fluency) over CF (linguistic knowledge and processing skill).

Data analysis shows that among the three aspects of UF; speed fluency (mean duration of syllables) is moderately related to linguistic knowledge and processing skill i.e. cognitive fluency. Results also showed moderate to weak correlation between one of the aspects (mean duration of silent pause) of breakdown fluency and CF. However, there was hardly any correlation between repair fluency and CF. Data of the present study provides new findings on the relation between L2 CF and UF since the result is somewhat different from the findings of previous research. Previous research found strong to moderate correlation between speed fluency (mean duration of syllables) and CF, and least correlation between mean duration of silent pause and CF (Baker_Smemoe, Dewey, Bown, & Martinsen, 2014; Kahng, 2014; De Jong, Steinel, Florijn, Schoonen, & Hulstijn, 2013), whereas the current results showed moderate correlation between speed fluency and CF, and some correlation between mean duration of syllables and CF. Thus, from the results of the current study it could be claimed that L2 cognitive fluency at the intermediate to lower-intermediate levels of L2 proficiency is not noticeable (no strong correlations) in terms of L2 utterance fluency. Speakers at this level of L2 proficiency may still be developing a stable and an efficient L2 system and therefore do not truly differ on these measures of L2 utterance fluency.

When compared with the results of the previous research, it can be claimed that certain features of L2 utterance fluency differ according to L2 proficiency level. The results of the present study suggest at intermediate to lower proficiency levels, employing L2 measures of utterance fluency alone for automatic scoring and assessment of L2 proficiency might not lead to sufficiently precise scoring. However, such L2 fluency measures could be useful for assessing at higher proficiency levels (Baker_Smemoe, 2014), as observed in the results of previous studies. Hence, for instructors aiming to find measures of L2 proficiency to identify learner progression (whether L2 learners have reached advanced level from intermediate level), automatic measures of L2 utterance fluency could provide sufficient evidence. Thus, aspects of utterance fluency that reflect learners' cognitive fluency could be taken into consideration while testing L2 fluency.

Baker_Smemoe, W., Dewey, D. P., Bown, J., Martinsen, R. A. (2014). Does Measuring L2 Utterance Fluency Equal Measuring Overall L2 Proficiency? Evidence from Five Languages. *Foreign Language Annals, 47*(4), 707-728. doi: 10.1111/flan.12110

De Bot, K. (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics, 13*, 1-24.

De Jong, N. H., Steinel, M. P., Florijn, A., Schoonen, R., & Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics, 34*(5), 893-916. doi:10.1017/S0142716412000069

Freed, B. F. (1995). Do students who study abroad become fluent? In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123–148). Amsterdam: John Benjamins.

Kahng, J. (2014). Exploring Utterance and Cognitive Fluency of L1 and L2 English Speakers: Temporal Measures and Stimulated Recall. *Language Learning 64*(4), 809–854.

Kormos, J., & D'enes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*, 145–164.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 3*, 387–417.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: Massachusetts Institute of Technology Press.

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes, 14*, 423–441.

Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *Canadian Modern Language Review, 65*, 395–412.

Segalowitz, N. (2010). *Cognitive Bases of Second Language Fluency.* New York: Routledge.

Van Moere, A. (2012). A psycholinguistic approach to oral language assessment. *Language Testing, 29*(3), 325-344.

Fluency and the use of foreign words in interviews with EFL learners

Sylvie DE COCK

Université catholique de Louvain, Université Saint-Louis Bruxelles

The Louvain International Database of Spoken English Interlanguage (LINDSEI) contains informal interviews with intermediate to advanced level learners of English as a foreign

language. The interviews follow the same set pattern and are made up of three main tasks: a personal narrative based on a set topic (an experience that taught them a lesson, a country that impressed them, or a film or play they liked/disliked), a free discussion mainly about university life, hobbies, foreign travel or plans for the future and a picture description. Although the interviews are all conducted in English, 'foreign' words, i.e. words from other languages than English, sometimes feature in both the interviewers' and the learner interviewees' contributions. Foreign words have been specially marked up in the LINDSEI corpus (<foreign> WORD(S) </foreign) and can therefore be retrieved automatically using WordSmith Tools for example.

In an investigation of the 'communication strategies' (Tarone 2005) used by the learners in half the Norwegian component of the LINDSEI corpus, Nacey and Graedler (2013: 352) discussed the use of foreign words ('code-switching') as an L1-based compensation strategy. They argued that, contrary to what other studies had shown, code-switching was a highly effective strategy in the interviews analysed as it contributed to 'a smooth flow of conversation'. This was however not taken any further. They only suggested that one of the reasons why code-switching was so effective was that the interviewers understood the learners' L1, i.e. Norwegian.

More recently, De Cock (2015) explored the use of foreign words in five of the subcorpora included on the LINDSEI CD-ROM (Gilquin et al. 2010), namely LINDSEI_Dutch, LINDSEI_French, LINDSEI_German, LINDSEI_Italian and LINDSEI_Spanish (each of these subcorpora contains between 140,000 and 80,000 words of interviewee and interviewer speech). The study reveals a rather complex picture of learners' use of foreign words. The foreign words, which come overwhelmingly but not exclusively from the learners' mother tongue, fall into four main categories: (1) lexical bridges, which help learners bridge vocabulary/lexical gaps (words/expressions that appear to be unknown or inaccessible to them; e.g. 'cotizar', 'des algues', 'lasser'),

(2) cultural/institutional bridges, which denote aspects of the education system, events, folklore, places, etc. typically associated with some of the regions/countries mentioned in the set topic and free discussion parts of the interviews (e.g. 'Tour de France', 'Parco Nazionale del Gran Paradiso', 'Vlaamse Opera', 'Abitur', 'gilles de Binche'),

(3) pragmatic/discourse bridges, which fulfil basic pragmatic/discourse functions in the learners' L1 (e.g. 'ja', 'allez', 'si', 'enfin', 'bueno'),

(4) foreign words used in direct speech reporting or in metalinguistic discussions (e.g. 'all she could say was <foreign> ich liebe dich </foreign>' - LINDSEI_DU , 'in Spanish they they call it <foreign> chela </foreign>' - LINDSEI_SP).

The analysis shows that not all the foreign words investigated could actually be labelled as 'communication strategies' defined as ‘systematic technique[s] employed by a speaker to express his meaning when faced with some difficulty’ (Corder 1981: 103). Pragmatic/discourse bridges are a case in point as their use is largely spontaneous and unintentional (e.g. 'because we are (er) two: <foreign> enfin </foreign> we we are three: children in my family and (er) two of us . are studying here so (er)' LINDSEI_FR).

This paper sets out to turn a 'fluency' spotlight on the use of foreign words in LINDSEI_Dutch, LINDSEI_French, LINDSEI_German and LINDSEI_Italian. The focus is on these four subcorpora as foreign words occur in at least three quarters of the interviews included in the components and are thus fairly evenly distributed. It is noteworthy that the interviewers

in the subcorpora on the LINDSEI CD-ROM either share the learners' L1 (e.g. LINDSEI_Dutch and LINDSEI_Italian) or are native speakers of English with at least some knowledge of the learners' L1 (e.g. LINDSEI_French and LINDSEI_German), which might arguably have some impact on the successful use of foreign words in the interaction (cf. Nacey and Graedler above). The aim is to examine whether or not and to what extent the foreign words used as lexical bridges, cultural/institutional bridges and pragmatic/discourse bridges could be labelled as (dis)fluency devices in the informal interviews with EFL learners under investigation. In other words, can the foreign words under study be seen to contribute (or not) to the smooth flow of the interviews?

Central to the notion of fluency used here are real-time pressure/processing and interaction management constraints (Rühlemann 2006). The various elements that typically co-occur with each of the three types of bridges are highlighted (e.g. 'I don't know how you say it in English', 'sort of/kind of', filled and unfilled pauses tend to co-occur with lexical bridges) as are turn positions and interviewers' reactions. The notion of fluency in interaction and fluent meaning co-construction (André & Tyne 2012) is also explored in the specific context of informal interviews, which do not share two of Clark's (1996) typical features of face-to-face conversation, namely self-determination (in informal interviews the turn-taking system is prespecified, Lazareton 1992) and self-expression (the interviewer has the right and obligation to ask questions and the interviewee has the obligation to answer these questions and to keep talking, Fiksdal 1990).

André, V. & Tyne, H. (2012) Compétence sociolinguistique et dysfluence en L2 In Kamber, A. & Skuipen-Dekens (eds) *Recherches récentes en FLE*. Bern: Peter Lang, 21-46. Clark, H. H. (1996) *Using Language*. Cambridge: Cambridge University Press.

Corder, S. P. (1981) *Error analysis and interlanguage*. Oxford: Oxford University Press.

De Cock, S. (2015) *An exploration of the use of foreign words in interviews with EFL learners: a(n) (effective) communication strategy?* Paper presented at LCR 2015, Nijmegen September 2015.

Fiksdal, S. (1990) *The Right Time and Pace: A Microanalysis of Cross-cultural Gatekeeping Interviews*. New Jersey: Ablex Norwood.

Gilquin, G., De Cock, S. & Granger, S. (eds) (2010) *The Louvain International Database of Spoken English Interlanguage. Handbook and CD-ROM*. Louvain-la-Neuve: Presses universitaires de Louvain.

Lazaraton, A. (1992) The Structural Organization of a language Interview: A Conversation Analytic Perspective. *System* 20/3, 373-386.

Nacey, S. & Graedler, A.-L. (2013) Communication strategies used by Norwegian students of English. In Granger, S., Gilquin, G. and Meunier, F. (eds) *Twenty Years of Learner Corpus Research: Looking Back, Moving Ahead*. Louvain-la-Neuve: Presses Universitaires de Louvain, 345-356.

Rühlemann, C. (2006) Coming to terms with conversational grammar: 'Dislocation' and 'dysfluency'. *International Journal of Corpus Linguistics*, 11 (4), 385–409.

Tarone, E. (2005) Speaking in a second language. In Hinkel, F. (ed.) *Handbook of Research in Second Language Teaching and Learning*. London: Routledge, 485-502.

Profiling French learners' productive (dis)fluency

Amandine DUMONT

Université catholique de Louvain

Over the past few decades, the notion of fluency has generated growing interest in the field of second and foreign language acquisition. In several recent studies (e.g. Bosker et al. 2013; Cucchiarini, van Doremalen & Strik 2010; Gilquin & Granger 2015), researchers have analysed the way learners of different proficiency levels or mother-tongue backgrounds make use of a number of devices such as filled and unfilled pauses, self-corrections or other speech management strategies. Despite the new insights that have emerged, however, there is still a lack of agreement among researchers regarding the scope of fluency: while a stream of research restricts the construct to the temporal aspects of speech (speech and articulation rate, pausal phenomena, etc.) (e.g. Ginther, Dimova & Yang 2010; Little et al. 2013; Zellner 1994), other conceptualisations are more far-reaching and encompass elements such as reformulations, false starts, discourse markers, formulaic sequences or pronunciation (e.g. Beliao & Lacheret 2013; Derwing, Thomson & Munro 2006; House 1996). In an attempt to get better insights into this kaleidoscope of components, a number of typologies have been put forward in the literature. Skehan and Tavakoli (Skehan 2003; Tavakoli & Skehan 2005; Tavakoli 2016), for instance, differentiate three subconstructs of fluency, namely speed, breakdown, and repair fluency. They argue that, while speed fluency relates to the speed of delivery and can be evaluated with speech rate measurements, breakdown fluency is concerned with the extent to which the speech flow is interrupted by pausal phenomena, and repair fluency includes phenomena that have to do with self-correction such as reformulations, replacements, false starts and repetitions.

Notwithstanding the absence of consensus on the precise definition of fluency, it is increasingly accepted in the literature that fluency results from the conjunction of a number of quantifiable and qualifiable variables and is consequently better analysed as a bundle of features. To date however, many studies have been restricted to the investigation of one of those features at a time without considering its potential interaction with other fluency features, and even fewer studies have considered performance variations within a (seemingly) homogenous dataset. One notable exception is Götz (2013), who analysed a comprehensive set of features and attempted to delineate the "fluency profiles" of German learners of English. In her study, she showed that speakers do have individual preferences for some features and that it is possible to distinguish speaker groups that have different ways of achieving fluency.

Against this backdrop, this paper investigates the separate contributions of a wide range of fluency features and examines how these correlate and interact with one another in the interlanguage of French learners of English. The features under investigation include the following: filled and unfilled pauses, restarts, false starts, repetitions, discourse markers and connectors, truncations, vowel lengthenings and foreign words as well as speech rate and mean length of pauses. The objectives are twofold. First, the analysis of the relationship between fluency features aims to investigate whether empirical corpus findings on fluency features support the taxonomy put forward by Skehan and Tavakoli, and more particularly the distinction between breakdown and repair fluency. In this respect, learner and native speaker data are also compared and contrasted in order to identify potential points of divergence. Second, as in Götz (2013) for German learners, the analysis seeks to reveal whether there are

different fluency profiles corresponding to different speaker types among a comparable group of French learners.

The study is based on the French component of the *Louvain International Database of Spoken English Interlanguage* (LINDSEI; Gilquin, De Cock & Granger 2010), which is a collection of 50 interviews of Belgian French-speaking university students of English as a foreign language of high-intermediate proficiency level. The native speaker data comes from LINDSEI's native speaker counterpart LOCNEC (the *Louvain Corpus of Native English Conversation*; De Cock 2004) that, likewise, contains interviews with 50 British English undergraduate native speakers. In both corpora, each interview contains three speaking tasks, namely a warming-up activity on a set topic, a free discussion and a picture description task, totalling c. 10 hours of learner and native language. The corpora have been time-aligned (which allows for the precise measurement of temporal phenomena) and subsequently annotated with the EXMARaLDA tool (Schmidt & Wörner 2014) for a wide number of (dis)fluency features such as filled and unfilled pauses, truncations, repairs, false starts or repetitions.

Results so far indicate that there are statistically significant, mostly moderate, correlations between a number of fluency features. These tend to support, although not perfectly, the distinction between breakdown and repair fluency. False starts, restarts, truncations and repetitions, for example, appear to be mutually related, while filled and unfilled pauses seem to form a separate group. The relationship between the two types of pauses however differs depending on the speaker group: whereas they are highly significantly correlated ($r = .45$; $p < .005$) in the native speaker data, it is not the case in LINDSEIFR. The analysis also reveals that the relationship between the two aspects of fluency, i.e. breakdown and repair, might be more complex than commonly assumed: in the data, filled pauses indeed correlate either negatively with false starts ($r = -.3$; $p < .05$; LINDSEI-FR) or positively with repetitions ($r = .4$; $p < .005$; LOCNEC) – and this finding also highlights an interesting difference between learner and native speaker fluency behaviour. Preliminary results from a hierarchical cluster analysis further show that not all the learners from the dataset behave similarly and that it is possible to identify different clusters among the learners' fluency performances.

Beliao, Julie & Anne Lacheret. 2013. Disfluency and discursive markers: when prosody and syntax plan discourse. *DiSS 2013: The 6th Workshop on Disfluency in Spontaneous Speech*, vol. 54, 5–9.

Bosker, Hans Rutger, Anne-France Pinget, Hugo Quené, Ted Sanders & Nivja De Jong. 2013. What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing* 30(2). 159–175.

Cucchiarini, Catia, Joost van Doremalen & Helmer Strik. 2010. Fluency in non-native read and spontaneous speech. *DiSS-LPSS*, 15–18.

De Cock, Sylvie. 2004. Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literature (BELL)* 2. 225–246.

Derwing, Tracey M., Ron I. Thomson & Murray J. Munro. 2006. English pronunciation and fluency development in Mandarin and Slavic speakers. *System* 34(2). 183–193.

Gilquin, Gaëtanelle, Sylvie De Cock & Sylviane Granger (eds.). 2010. *LINDSEI. Louvain International Database of Spoken English Interlanguage*. Presses Universitaires de Louvain. Louvain-laNeuve.

Gilquin, Gaëtanelle & Sylviane Granger. 2015. Learner language. In Douglas Biber & Randi Reppen (eds.), *The Cambridge Handbook of Corpus Linguistics*, 418–435. Cambridge: Cambridge University Press.

Ginther, April, Slobodanka Dimova & Rui Yang. 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.

Götz, Sandra. 2013. *Fluency in Native and Nonnative English Speech*. (Studies in Corpus Linguistics (SCL) 53). Amsterdam & Philadelphia: John Benjamins Publishing Company.

House, Juliane. 1996. Developing Pragmatic Fluency in English as a Foreign Language: Routines and Metapragmatic Awareness. *Studies in Second Language Acquisition* 18(2). 225–252.

Little, Daniel R., Raoul Oehmen, John Dunn, Kathryn Hird & Kim Kirsner. 2013. Fluency Profiling System: An automated system for analyzing the temporal properties of speech. *Behavior Research Methods* 45(1). 191–202.

Schmidt, Thomas & Kai Wörner. 2014. EXMARaLDA. In Jacques Durand, Ulrike Gut & Gjert

Kristoffersen (eds.), *The Oxford Handbook of Corpus Phonology*, 402–419. Oxford University Press.

Skehan, Peter. 2003. Task-based instruction. *Language Teaching* 36(1). 1–14.

Tavakoli, Parvaneh. 2016. Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching* 54(2). 133–150.

Tavakoli, Parvaneh & Peter Skehan. 2005. Strategic planning, task structure, and performance testing. In Ellis Rod (ed.), *Planning and task performance in a second language*, 239–273.

Zellner, Brigitte. 1994. Pauses and the temporal structure of speech. In E. Keller (ed.), *Fundamentals of speech synthesis and speech recognition*, 41–62. Chichester: John Wiley.

Temporal characteristics of speech during fluency-inducing conditions in stuttering

Simone FALK[1, 2], Ramona SCHREIER[2], Elena MASLOW[2], Georg THUM[2], Phil HOOLE[2]

CNRS/Université Paris 3 - Sorbonne Nouvelle[1], Ludwig-Maximilians-Universität München[2]

Stuttering is a developmental fluency disorder provoking interruptions of the rhythmic flow of speech. Around 5-9 % of children and adolescents are, at least temporarily, affected by stuttering (Yairi & Ambrose, 2014). Disfluencies which are characteristic of stuttering are repetitions of segments and syllables, blockades, and segmental prolongations. Moreover, breathing difficulties can lead to untimely interruptions within words or phrases. However, stuttering has been shown to reduce considerably during so-called fluency-inducing conditions, such as singing or speaking along with a metronome (Andrews et al., 1982; Wingate, 1969). The aim of the present contribution is to examine temporal characteristics (i.e., variability, timing; see e.g., Janssen & Wieneke, 1987; Onslow et al, 1992) of perceptually fluent speech in individuals who stutter during two fluency-inducing conditions. In two studies, we tested groups of German-speaking children and adolescents who do and do not stutter while singing and while reading a word list with a metronome on the temporal variability and timing of their segmental productions. In the first study on singing (Falk, Maslow, Thum & Hoole, 2016), we found that 16 adolescents (11-16 years old) who do and do not stutter reduced Voice Onset Times (VOTs) during singing compared to speaking, but only adolescents who stutter also reduced the temporal variability of VOT during singing vs. speaking. In another study (Falk, Schreier & Thum, in preparation), we tested 28 children and adolescents (8-17 years old) on

their ability to align monosyllabic words to the pace of a metronome. Results showed that children and adolescents who stutter systematically delayed the start of vocalic nuclei in relation to the metronome beat, compared to children and adolescents who do not stutter. Results are discussed in light of related studies on fluency-reducing techniques in stuttering (Davidow et al., 2011; Stager et al., 2003) and recent theories about rhythmic predictions and predictive timing (e.g., Kotz & Schwartze, 2016; Maes et al., 2014) fostering the coupling between perception and action and thereby, the fluent production of speech.

Andrews, G., Howie, P. M., Dozsa, M., & Guitar, B. E. (1982). Stuttering: speech pattern characteristics under fluency-inducing conditions. *Journal of Speech and Hearing Research*, 25(2), 208–216

Davidow, J. H., Bothe, A. K., & Ye, J. (2011). Systematic studies of modified vocalization: speech production changes during a variation of metronomic speech in persons who do and do not stutter. *Journal of Fluency Disorders*, 36(2), 93–109.

Falk, S., Maslow, E., Thum, G., & Hoole, P. (2016). Temporal variability in sung productions of adolescents who stutter. *Journal of Communication Disorders*, 62, 101–114.

Janssen, P., & Wieneke, G. (1987). The effects of fluency inducing conditions on the variability in the duration of laryngeal movements during stutterers' fluent speech. In H. F. M. Peters, & W. Hulstijn (Eds.), *Speech motor dynamics in stuttering* (pp. 337–344). New York, NY: Springer

Kotz, S. A., & Schwartze, M. (2016). Motor-timing and sequencing in speech production: a generalpurpose framework. In G. Hickok, & S. Small (Eds.), *Neurobiology of language* (pp. 717–724).Amsterdam a.o: Academic Press.

Maes, P.-J., Leman, M., Palmer, C., & Wanderley, M. M. (2014). Action-based effects on music perception. *Frontiers in Psychology*, 4, 1008.

Onslow, M., van Doorn, J., & Newman, D. (1992). Variability of acoustic segment durations after prolonged-speech treatment for stuttering. *Journal of Speech Language and Hearing Research*, 35(3), 529.

Stager, S. V., Jeffries, K. J., & Braun, A. R. (2003). Common features of fluency-evoking conditions studied in stuttering subjects and controls: an H215O PET study. *Journal of Fluency Disorders*, 28(4), 319–336.

Wingate, M. E. (1969). Sound and pattern in artificial fluency. *Journal of Speech and Hearing Research*, 12(4), 677–686.

Yairi, E., and Ambrose, N. G. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders*, 38, 66–87.

Between pragmatic function and bad reputation: Partner models mediate the use and interpretation of hesitation markers

Kerstin FISCHER & Nathalie SCHÜMCHEN

University of Southern Denmark

Folk notions of hesitation markers (also called 'filled pauses') suggest that their occurrence is related to cognitive load, for instance, nervousness, lack of knowledge, or lack of concentration (Fischer, 2000; Fox Tree, 2001; Fraundorf & Watson 2014). Correspondingly, previous work on hesitation markers has attempted to correlate their occurrence with utterance planning (Boomer, 1965; Clark, 2006; Corley & Stewart, 2008), lexical choice (Suh, 2000) and cognitive processing in general (Christenfeld, 1994; Lindsey, Greene, Parker, & Sassi, 1995). However,

these attempts have not been entirely successful; while increased cognitive load is correlated with increasing numbers of disfluencies (e.g. Su & Luz 2016), not all of their occurrences can be explained in this way (e.g. Levinson 1983; Schegloff 2010). More recently, numerous pragmatic functions of hesitation markers have been suggested both from conversation analytic and psychological perspectives (e.g. Clark & Fox Tree, 2002; Fischer, 2000; Schegloff, 2010). The discrepancy is not yet fully reconciled: The question remains how it is possible for hesitation markers to have such a bad reputation when they may fulfill so many useful functions.

Building on the model for addressee orientation developed in Fischer (2016), we argue that hesitation marker use is crucially influenced by speakers' partner models, and that, conversely, the interpretation of hesitation markers crucially depends on listeners' models of the respective speaker. In particular, we suggest that the fact that hesitation markers are in systematic opposition to pauses indicates pragmatic function and hence that these markers are signs (see also Clark & Fox Tree, 2002). The function of these signs is to indicate ongoing thought processes (Fischer, 2000), not as symptoms of cognitive processes but as socially governed signs that indicate such processes. For instance, the indication of ongoing thought is pragmatically required before rejecting an invitation; for example, responding 'no' without delay to the question 'how about dinner tonight?' is more rude than marking the same answer as hesitant to indicate that it was a hard decision (Levinson, 1983:333). In other words, signaling ongoing thought processes should at least have the following pragmatic functions:[1] o Hesitation markers indicate chunks of information and important words and hence help structure the information;

- making one's thought processes transparent/accessible to the partner should influence likeability;
- hesitation markers indicate ad hoc production in comparison to canned, prefabricated speech, which, in turn, serves a socialinteractive function (similar to repair in conversation, cf. Schegloff et al. 1977; Jefferson 1972).

Obviously, this does not mean that they are strategically used, conscious signals; like most other grammatical choices we assume the use of hesitation markers to be functional, yet below levels of conscious awareness.

To test these hypotheses, we designed a between-subject experiment in which participants were presented with three variations of six original stimuli containing the hesitation marker 'uh' from six different TED talks presented by six different speakers. The hesitation markers occur in different positions and hence different functions, namely marking topic boundaries (Luz & Su, 2010) and important words (Fox Tree, 2001). In addition to the original version of the stimulus (utterance including 'uh'), the six stimuli were manipulated in two ways. First, the hesitation marker was edited out, and second, the hesitation marker was edited out and replaced by silence. This means that each of our participants was presented with six stimuli, each uttered by a different speaker: two original stimuli including 'uh', two without hesitation, and two with silence instead of 'uh'. Each utterance was then followed by a comprehension question about a detail of the utterance. Moreover, in order to test our hypothesis that hesitation markers serve an important function regarding addressee orientation, we framed the questionnaires in one of two ways: We told participants that they were going to listen to excerpts either from 'great speakers' or from 'great teachers'. This framing was followed by a list of attributes (e.g. 'intelligent', 'friendly', 'knowledge about topic') that the participants had to rank according to

their expectations in accordance with the framing. Thus, two differently framed questionnaires that each included two utterances of each condition in random order were designed (resulting in 90 possible combinations) and distributed to 200 native speakers of English. The dependent measures were participants' responses to questions about pragmatic function, but also about their predispositions and partner model. The first set of questions therefore addressed to which extent the speaker is perceived as trying to get something across, is involved, wishes the listener to really understand, and is perceived as friendly, likeable and polite (among other categories). The second set of questions concerned participants' expectations about good speakers and good teachers.

A first analysis of the data shows significantly more positive ratings for utterances with hesitation markers marking important words as well as more correct answers to the comprehension questions. Furthermore, the speakers that prefaced important words with 'uh' were rated as significantly less nervous and unconcentrated. With regard to the framings, the results show significant differences in the order of prioritized attributes. 35.8% of the participants who heard 'great speakers' put 'speaks fluently' on the first rank while those expecting teachers put highest priority on 'knowledge about topic' (45.5%). In both cases, 'high education' was regarded as least important (74.4% in condition 1; 48.8% in condition 2). This indicates that expectations regarding 'fluency' differ across situations. While further analysis will investigate the correlation between the framing and participants' judgments of speech with and without hesitation markers in more detail, the results so far suggest that hesitation markers indeed fulfill important information structuring and interpersonal functions and thus do not deserve their bad reputation.

Boomer, D. (1965). Hesitation and Grammatical Encoding. *Language and Speech, 8*, 148-158.

Christenfeld, N. (1994). Options and Ums. *Journal of Language and Social Psychology, 13*(2), 192-199.

Clark, H. H. (2006). Pauses and hesitations: Psycholinguistic approach *Encyclopedia of Language and Linguistics 2nd Edition*. Oxford: Elsevier.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*(1), 73-111.

Corley, M., & Stewart, O. W. (2008). Hesitation Disfluencies in Spontaneous Speech: The Meaning of um. *Language and Linguistics Compass, 2*(4), 589602.

Fischer, K. (2000). *From cognitive semantics to lexical pragmatics: The functional polysemy of discourse particles*: Walter de Gruyter.

Fischer, K. (2016). *Designing Speech for a Recipient: The roles of partner modeling, alignment and feedback in so-called 'simplified registers'* (Vol. 270). Amsterdam, Philadelphia: John Benjamins Publishing Company.

Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory & Cognition, 29*(2), 320-326.

Fraundorf, Scott H. and Watson, Duane G. (2014). Alice's adventures in umderland: psycholinguistic sources of variation in disfluency production. *Language, Cognition and Neuroscience* 29, 9, 1083-1096.

Su, Jing and Luz, Saturnino (2016). Predicting Cognitive Load Levels from Speech Data. *Recent Advances in Nonlinear Speech Processing*. Smart Innovation, Systems and Technologies 48, pp. 255-263

Levinson, S. C. (1983). Pragmatics (Cambridge textbooks in linguistics).

Lindsey, A. E., Greene, J. O., Parker, R. G., & Sassi, M. (1995). Effects of advance message formulation on message encoding: Evidence of cognitively based hesitation in the production of multiple-goal messages. *Communication Quarterly, 43*(3), 320-331.

Luz, S., & Su, J. (2010). *The relevance of timing, pauses and overlaps in dialogues: detecting topic changes in scenario based meetings.* In Proceedings of INTERSPEECH 2010 (Chiba, Japan, 2010), ISCA, pp. 1369–1372.

Schegloff, E. A. (2010). Some Other "Uh(m)"s. *Discourse Processes, 47*(2), 130174.

Suh, K.-H. (2000). Distal demonstratives as fillers. *Language Research, 36*(4), 887-903.

Repeats in native and learner English

Tomáš GRÁF[1], Lan-fen HUANG[2]

Charles University[1], Shih Chien University[2]

Spontaneous speech production is notoriously interspersed with different types of disfluencies of which repeats are some of the most frequent ones. This applies to both native speakers and language learners. But do they produce repeats in a similar way? In our study we have compared recordings of 50 native speakers of English, 50 Czech and 50 Taiwanese advanced learners of English, and tried to answer questions regarding the frequency of repeats, the types of repeated words and their individual frequencies, the length of the repeated segments, between-speaker variability within the individual groups, the nature of any established differences thereof, and correlations with other fluency variables as well as with the learners' proficiency.

The data for our analysis has been derived from the spoken parallel corpora LINDSEI (namely its Czech and Taiwanese subcorpora) and LOCNEC. Each of the 150 interviews is approximately 15 minutes long. The total number of tokens is approximately 300,000. The recordings have been orthographically transcribed, and all instances of repeats have been identified and tagged using a semi-automatic computer script which the author of the study developed for the purpose. Any instances of repetitions which had a semantic function (e.g. intensification) were removed. The tagging system we designed enables differentiation between the length of repeats (e.g. one- or two-word repeats), numbers of repetitions, the word classes and discourse functions. The tagged corpora were then processed using a concordancer which facilitated the quantification of results and the sorting into categories for deeper analyses.

The three corpora contain 5,253 instances of repeats, 77% of which are presented by oneword repeats, 19% by two-word repeats, and 4% by repeated stretches of three or more words. Whilst these results are almost identical for the native and non-native corpora, differences appear in the mean rate of frequency, where the native speakers produce 1.5 (SD=0.87) repeats per hundred words (phw), the Czech learners 1.9 (SD=1.18) repeats phw, and the Taiwanese learners 2.15 (SD=1.47) repeats phw. As regards betweenspeaker variability, this is greater in the Czech and Taiwanese subcorpora which have a larger number of speakers with a higher repeat rate than the mean repeat rate observed in the respective subcorpora. This might indicate that the learners feel a greater need to use repeats to maintain fluency. It remains to be established whether this finding can be correlated to proficiency (the corpus is currently being rated for proficiency) and establish whether less proficient users show a greater reliance on repeats or whether, on the contrary, more proficient learners use repeats in a more native-like manner.

As regards the analysis of the most frequently repeated word classes, all speakers repeat especially pronouns, prepositions, articles and conjunctions. The repeated pronouns consist mainly of subject pronouns (especially *I, we, he, they*) and possessives (esp. *my*). The most frequently repeated conjunctions are *and, but* and *if*. The most frequently repeated prepositions include *in, of, for, to* and *about*. The learners significantly ($p < 0.001$) overuse repeats of practically all word classes except the indefinite article and contracted forms. The most underused repeats of contractions are *I've, I'll, I'd, he's* and *it's*.

No correlations have been found between the use of repeats and other fluency variables. These include speech rate, and the frequency of filled and unfilled pauses. There does not appear to be any correlation between the learners' accuracy (operationalised as the rate of grammatical, lexical and lexico-grammatical errors) and their use of repeats. A further examination of those learners who have an especially high repeat rate has revealed no typical pattern regarding the measures of these phenomena. We have yet to examine possible correlation with proficiency, which in the Czech subcorpus is likely to range between the CEFR levels B2 and C2, and in the Taiwanese subcorpus between B1 and C1 (based on preliminary ratings).

The results show that the learners had internalised the native strategy of producing repeats of primarily function words at the beginning of utterances, clauses or constituents as described in Maclay & Osgood (1959), Clark and Wasow (1998), Biber et al. (1999) and Kjellmer (2008). The results may be compared to Götz's (2013) study of the same phenomena with the German subcorpus of LINDSEI. As was true for the our learners, Götz also found a significant underuse of repeats of articles and contractions by the learners. However, she found an overall significant underuse of repeats by the German learners which contrasts with our finding of the Czech and Taiwanese learners overusing repeats. A close inspection of some of the individual speakers revealed idiosyncrasies in the use of repeats and considerable between-speaker variation. This might be an indication of the potential of the phenomenon for forensic science.

The comparison of the three corpora has revealed that the learners have successfully adopted a frequent native strategy of producing repeats to buy time for planning speech and for resolving arising problems. However, they mostly produce these with a higher frequency (except contractions and indefinite articles). Further work is to be carried out investigating possible correlations between our findings and the learners' proficiency.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow, England; New York: Pearson Education ESL.

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

Götz, S. (2013). *Fluency in Native and Nonnative English Speech*. Amsterdam: John Benjamins Publishing Company.

Kjellmer, G. (2008). Self-repetition in spoken English discourse. In T. Nevalainen & I. Taavitsainen (Eds.), *The dynamics of linguistic variation: corpus evidence on English past and present ; [... a selection of articles based on papers presented at the 27th Conference of the International Computer Archive of Modern and Medieval English (ICAME) in Helsinki in May 2006]*. Amsterdam: John Benjamins Publishing Company.

Maclay, H., & Osgood, C. E. (1959). Hesitation Phenomena in Spontaneous English Speech. WORD, *15*(1), 19–44.

McDougall, K., Duckworth, M. & Hudson, T. (2015). Individual and Group Variation Dysfluency Features: A Cross-Accent Investigation. *Proceedings of the 18th International Congress of Phonetic Science (ICPhS 2015).* Glasgow: University of Glasgow.

Prosodic variation of identical repetitions as a function of their properties and editing terms: A large-scale corpus study on French speech

Iulia GROSMAN, George CHRISTODOULIDES, Liesbeth DEGAND, Anne-Catherine SIMON

Université catholique de Louvain

Lexical *Identical Repetitions,* where the speaker produces the same lexical form multiple times in a row, constitute a substantial part of speech production (e.g. *<the the> small thing* ; *<the thing the thing> is that*). Several typologies and potential function(s) have been proposed for these repetitions, depending on the research paradigm: cognitive, textual, stylistic, argumentative, conversational, interactional, sociolinguistic, etc. (Bazzanella 2011: 252). Previous linguistic studies usually focus on only one pattern of syntagmatic and/or acoustic realisation of repetition, or on one function, almost exclusively (cf. Hieke 1981; Shriberg 1995).

A survey of the literature reveals several ways to model the phenomenon. Identical repetitions, along with several other types of structured disfluencies, can be described as a sequence of three contiguous regions:

(reparandum) * interruption point (interregnum, including optional editing terms) (reparans)

The *reparandum* is the part of the utterance that is repeated. The interregnum is the region between the reparandum and the repair. It may optionally include explicit editing terms, i.e. words or phrases used by the speaker to signal the correction (e.g. discourse markers). The repair is the continuation of the message that follows the disfluency, so that if the first two regions are removed the remainder is lexically fluent. The interruption point is the point between the *reparandum* and the *interregnum*: this instance in time does not necessarily coincide with the moment the speaker detected the trouble or with his intention to alter the utterance (Shriberg, 2001).

On the morpho-syntactical level, there is general agreement that identical repetitions tend to occur on monosyllabic function words, and especially on articles and pronouns (e.g. Dister 2007; Candea 2000; Henry, Campione & Véronis, 2004). On the syntactic level, similar to other types of (dis)fluencies (Levelt 1983), identical repetitions present regularity in their structure, and they tend to co-occur with silent and filled pauses and editing terms. Acoustically, it has been shown that segments in the reparandum and the interregnum are lengthened relatively to the reparans, or compared to occurrences of the same units in fluent contexts (Shriberg 1999). Finally, prosodically, some studies have shown that the pitch of the onset (beginning) of the reparans is higher to the pitch at the offset (end) of the reparandum (Savova & Backenko 2003). In summary, the effects of identical repetitions have been studied on three levels:

While previous studies have focused separately on syntagmatic, morpho-syntactic, or acoustic properties of identical repetition (e.g. left and right periphery, types of interruption point (Shriberg 1995), type of tokens (Candea 2000, Clark & Wasow 1998), pitch and duration effect in reparandum, editing terms and reparans), we aim here to take into account all three levels of analysis to empirically model the phenomenon of identical repetition[2] across speaking styles. To this end we compiled five phonetically aligned corpora: the LOCAS-F corpus (Martin, Degand, and Simon 2014), the C-Humour corpus (Grosman 2016), the Driving Simulator Cognitive Load corpus (Christodoulides 2016), the CPhonogenre corpus (Prsir, Goldman, and Auchlin 2014), and the Rhapsodie corpus (Lacheret et al. 2014). The compilation covers 31 speaking styles, includes a total of 276 different samples, and its total duration is 17,4 hours (186.895 tokens). Identical repetitions have been automatically detected using DisMo (Christodoulides, Avanzi, and Goldman 2014) and manually verified using ARCFluency Disfluency Annotation Scheme (Crible et al. 2016). Approximately, 3000 repetition sequences have been extracted.

Our study pursues several goals. Firstly, we try to bring an empirical response to the question of typology, i.e. the categorisation of identical repetitions based on their syntactic structure. We have described each repetition based on: (1) the number of tokens repeated, (2) the number of repetitions of each string of tokens, (3) the presence/absence and the type of interregnum, with the (a) presence of a silent pause and/or (b) of editing term(s), and finally, (4) the left and right immediate context which may be (dis)fluent (including silent and filled pauses, discourse markers, truncations, etc.).

Figure 1 gives insight into the association between factors; it also guides the selection of a set of useful features for the acoustic description of different repetition patterns. Overall, 86% of repetitions are not followed by a (silent or filled) pause, and 88% do not include an interregnum; although there is no significant association between the presence of a pause following the repetition and the presence of an interregnum ($\chi^2 = 2.63$, df = 1 p = 0.10).

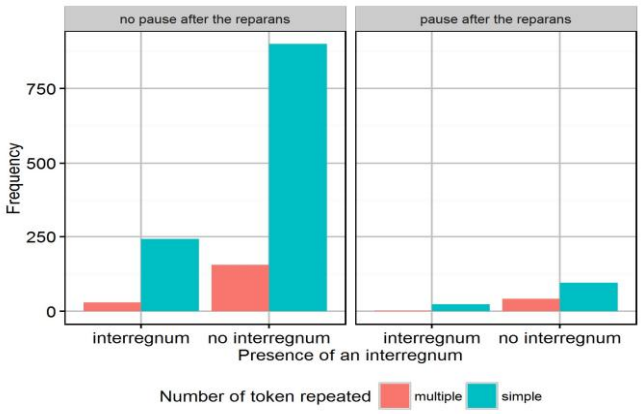


Figure 1 – Relative frequency of types of repetitions in 5 corpora.

2 In order to avoid pauses due to the management of interaction ("gaps") rather than solely to discourse processing (Heldner and Edlund 2010), we only included in the study *reparandum, editing terms* and *reparans* occurring within one single speaker speech turn.

The second part of the analysis focuses on the relationship between repetition types and immediate contexts, on the one hand, and their prosodic properties of the other hand. With respect to duration, our results indicate that in the case of simple repetitions, those with a pause inside the interregum tend to have a relatively longer reparans (decrease in local articulation rate) compared to repetitions where the reparandum is not followed by a pause. This confirms Shriberg's (1995) observations on English (Figure 2). With respect to intonation, we have performed mixed-effect modelling (with random effects on speaker) in order to compare the difference between the offset and onset of reparandum and reparans for each repetition pattern observed in the corpus. Different techniques are necessary to model the prosodic properties of monosyllabic and polysyllabic repetitions: while in the former we can readily compare the pitch contour of the reparandum and the reparans at the syllable level, the latter case requires comparisons across stylised pitch contours extending to the lexical level.

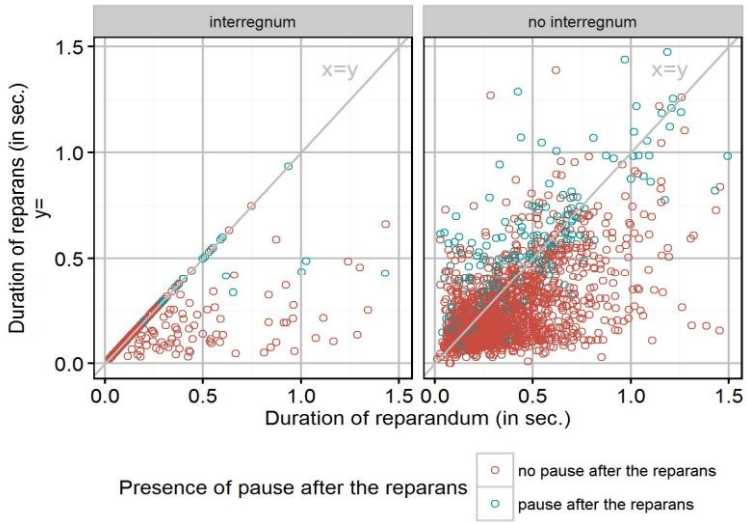


Figure 2 – Relative duration of segments reparans and reparandum in regard with structure of the repetitions

Bazzanella, Carla. 2011. 'Redundancy, Repetition, and Intensity in Discourse'. *Language Sciences* 33 (2): 243–54. doi:10.1016/j.langsci.2010.10.002.

Candea, Maria. 2000. 'Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané. Etude sur un corpus de récits en classe de français.' Thèse de doctorat, Université de la Sorbonne nouvelle - Paris III. https://tel.archivesouvertes.fr/tel-00290143/document.

Christodoulides, George. 2016. 'Effect of Cognitive Load on Speech Production and Perception'. PhD Thesis, Université catholique de Louvain.

Christodoulides, George, Mathieu Avanzi, and Jean-Philippe Goldman. 2014. 'DisMo: A Morphosyntactic, Disfluency and Multi-Word Unit Annotator. An Evaluation on a Corpus of French Spontaneous and Read Speech'. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC) 2014*, 3902–7. Reykjavik, Iceland: European Language Resources Association (ELRA). http://www.lrecconf.org/proceedings/lrec2014/pdf/929_Paper.pdf.

Clark, Herbert H., and Thomas Wasow. 1998. 'Repeating Words in Spontaneous Speech'. *Cognitive Psychology* 37 (3): 201–42. doi:10.1006/cogp.1998.0693.

Cole, Jennifer S., Mark Hasegawa-Johnson, Chilin Shih, Heejin Kim, Eun-Kyung Lee, Hsin-yi Lu, Yoonsook Mo, and Tae-Jin Yoon. 2005. 'Prosodic Parallelism as a Cue to Repetition and Hesitation Disfluency'. In *Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop*, 53–58. Aix-en-Provence, France.

Crible, Ludivine, Amandine Dumont, Iulia Grosman, and Ingrid Notarrigo. 2016. 'Annotation Manual of Fluency and Disfluency Markers in Multilingual, Multimodal, Native and Learner Corpora. Version 2.0.' Technical Report. Université catholique de Louvain and Université de Namur.: Université catholique de Louvain and Université de Namur.

Dister, Anne. 2007. 'De La Transcription À L'étiquetage Morphosyntaxique. Le Cas de La Banque de Données Textuelles Orales VALIBE'. Thèse de doctorat, Université catholique de Louvain.

Grosman, Iulia. 2016. 'How Do French Humorists Manage Their Persona across Situations? A Corpus Study on Their Prosodic Variation'. In *Metapragmatics of Humor: Current Research Trends*, edited by Leonor Ruiz-Gurillo, 149–77. JB IVITRA. John Benjamins Publishing Company.

Heldner, Mattias, and Jens Edlund. 2010. 'Pauses, Gaps and Overlaps in Conversations'. *Journal of Phonetics* 38 (4): 555–68. doi:10.1016/j.wocn.2010.08.002.

Henry, Sandrine, Estelle Campione, and Jean Véronis. 2004. 'Répétitions et Pauses (Silencieuses et Remplies) En Français Spontané'. In *Actes Des XXVè Journées d'Etude Sur La Parole (JEP'2004)*, 1–4. Fes, Maroc: ISCA-AFCP. http://sites.univ-provence.fr/~veronis/pdf/2004Henry-JEP.pdf.

Henry, Sandrine, and Berthille Pallaud. 2002. 'Étude Des Répétitions En Français Parlé Spontané Pour Les Technologies de La Parole'. In *Actes Des 4e Rencontres Des Étudiants Chercheurs En Informatique Pour Le Traitement Automatique Des Langue*, 467–476. Nancy, France: Association pour le Traitement Automatique des Langues. http://www.atala.org/taln_archives/RECITAL/RECITAL-2002/recital-2002-long-007.pdf.

———. 2003. 'Word Fragments and Repeats in Spontaneous Spoken French'. In *Proceedings of DiSS'03, Disfluency in Spontaneous Speech Workshop*, edited by Robert Eklund, 77–80. Gothenburg Papers in Theoretical Linguistics. http://www.iscaspeech.org/archive_open/diss_03/dis3_077.html.

Hieke, Adolf E. 1981. 'A Content-Processing View of Hesitation Phenomena'. *Language & Speech* 24 (2): 147–60.

Lacheret, Anne, Sylvain Kahane, Julie Beliao, Anne Dister, Kim Gerdes, Jean-Philippe Goldman, Nicolas Obin, Paola Pietrandrea, and Atanas Tchobanov. 2014. 'Rhapsodie: A ProsodicSyntactic Treebank for Spoken French'. In *Language Resources and Evaluation Conference*. http://hal.upmc.fr/hal-00968959/.

Levelt, W. J. 1983. 'Monitoring and Self-Repair in Speech'. *Cognition* 14 (1): 41–104.

Martin, Laurence J., Liesbeth Degand, and Anne-Catherine Simon. 2014. 'Forme et fonction de la périphérie gauche dans un corpus oral multigenres annoté'. *Corpus*, Eléments initiaux dans la phrase : approches inter-genres et inter-langues, , no. 13: 243–65.

Prsir, Tea, Jean-Philippe Goldman, and Antoine Auchlin. 2014. 'Prosodic Features of Situational Variation across Nine Speaking Styles in French'. *Journal of Speech Sciences* 4 (1): 41–60.

Savova, Guergana, and Joan Bachenko. 2003. 'Prosodic Features of Four Types of Disfluencies'. In *ISCA Tutorial and Research Workshop on Disfluency in Spontaneous Speech*. http://www.isca-speech.org/archive_open/diss_03/dis3_091.html.

Shriberg, Elizabeth. 1995. 'Acoustic Properties of Disfluent Repetitions'. In *Proceedings of the International Congress of Phonetic Sciences*, 4:384–387. http://purl.tue.nl/572383975006721.pdf.

———. 1999. 'Phonetic Consequences of Speech Disfluency'. In *Proceedings of the 14th ICPhS Conference*, 619–22. San Francisco: DTIC Document. http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA451506. _____. 2001. 'To "errrr" is human: ecology and acoustics of speech disfluencies'. *Journal of the International Phonetic Association* 31(1): 153-169.

An Investigation of stuttering in Persian-speaking children based on the CALMS Assessment (Cognitive, Affective, Linguistic, Motor and Social)

Yasaman JALILIAN[1], Fariba YADEGARI[1], Abbas EBADI[2]

University of Social Welfare and Rehabilitation Sciences[1], Baqiyatallah University of Medical Sciences[2]

Purpose: Stuttering is known as a multi factorial fluency disorder. CALMS assessment assesses stuttering from a multidimensional perspective, and valid and reliable Persian version of CALMS (CALMS-P) is available. Due to importance of stuttering treatment as a multi factorial disorder, this study involved the use of the CALMS-P in order to assess the impact of variables such as age, gender and family history on stuttering components in 115 Persianspeaking children who stutter. Method: the relation between the five components of CALMS-P with age, gender and family history of stuttering variables were examined. The data were analyzed using one-way ANOVA and independent t- test by SPSS software version 17. A p-value less than 0.05 were considered as statistically significant. A one-way ANOVA was performed on the stuttering factors at different age to see if there was a relationship between components and age of people who stutter. In the second analysis, independent- t test was performed on the stuttering factors at different gender and family history of stuttering. Results showed there is difference between girls and boys in linguistic and social components. The family history of stuttering had significant difference with affective, linguistic, motor and social components of stuttering. Results also tentatively suggest that people who stutter with increasing in age may show higher cognitive information of stuttering. Conclusion: These findings have implications for treatment such as the necessity to address the cognitive aspects of stuttering in different age and the need for additional clinical resources to be invested in stuttering treatment.

Validation of a multidimensional model of stuttering for Persian-speaking children who stutter

Yasaman JALILIAN[1], Fariba YADEGARI[1], Abbas EBADI[2], E. Charles HEALEY[3], Ebrahim HAJIZADEH[4]

University of Social Welfare and Rehabilitation Sciences[1], Baqiyatallah University of Medical Sciences[2], University of Nebraska[3], Tarbiat Modares University[4]

Purpose: The purpose of the present study was to validate a multidimensional model of stuttering usingcognitive, affective, linguistic, motor, and social (CALMS) assessment instrument for Persian-speaking children who stutter.

Method: The CALMS assessment instrument was translated to Persian (CALMS-P) using an IQOLA translation process. Content validity was examined by the content validity of individual items (I-CVI) and the overall scale (SCVI).The CALMS-P was administered with 115 Persian-Speaking children who stutter (age 7-14 years, 92 boys and 23 girls). Construct

validity was probed through confirmatory factor analysis (CFA) using LISREL 8.8 and internal consistency was analyzed by Cronbach's alpha via SPSS 17 software.

Results: CALMS-P was an 18-item Persian assessment instrument with five subscales measuring five dimensions of stuttering i.e. cognitive, affective, linguistic, motor and social components. The model provided a fair fit ($\chi 2$ /df = 1.60; RMSEA=0.07) and a good internal consistency (Cronbach's alpha =0.85) for the total instrument.

Conclusion: the CALMS-P assessment instrument which is based on the multidimensional model of stuttering proved to be a valid tool, useful for the therapists to assess and manage the stuttering of Persian-speaking children given the multi-factorial perspective of stuttering.

How do Deaf native signers perceive fluency in adult L2/M2 learners of German Sign Language?

Thomas KAUL, Alejandro OVIEDO, Leonid KLINNER, Reiner GRIEBEL & Andreas GLATZ

University of Cologne

Introduction

According to the Common European frame of Reference – CEFR[1], ‚fluency' in German Sign Language (DGS) is one of the main evaluation criteria for L2-language proficiency. However in research practice fluency in DGS is regarded as a desideratum, which makes it hard to teach and assess.This pilot study attempts to initiate a discussion about what fluency actually implies and how it can be assessed.

Even though there is not DGS equivalent for the word ‚fluency', native signers seem to be aware of its construct, as they are able to rate DGS learners' sign language production according to its fluency. A term equivalent to ‚fluent' is associated with native-like signed productions. Signing of late signers and late learners of DGS as L2 is related to terms equivalent to ‚non-fluent' [2].

In order to operationalize ‚fluency' and assess it in a more standardized way, we conducted a little survey among deaf signers, many of who were DGS teachers, too (53%). Our aim was to identify possible criteria of fluency, or to put in other words, what fluency consists of.

Methods

Two fellow Deaf researchers of our university, who also teach DGS to adult L2/M2 learners, selected two video samples from the Cologne Corpus of DGS as L2/M2 [3]. Each of the videos was about 2,3 minutes long and showing a monologue of a hearing adult L2/M2 learner of DGS retelling the same story. Both signers have successfully completed the CEFR B2 DGS-level and the videos were part of their course´s final evaluation. The student of **Video 1** received the best mark of the course, the student of **Video 2** a considerable lower one.

Video 1 was considered fluent; **Video 2**, less fluent. Criteria considered for the selection were based on two previous studies on sign language data:

a. amount and nature of pauses [4,5,6,7];

b. intelligibility –i.e. semantic coherence [5,6];

c. amount and frequence of disfluency markers (self-repairs, repetitions and meaningless gestures) [4,5,7];

d. grammatical correctness (grammatic errors understood as disfluency markers [6]);

A linguistic comparison of both videos evidenced that

- Student 1 made less and shorter pauses than student 2. Pauses made by student 1 were always filled (either with frozen signs or meaningless gestures) and systematically accompanied by non-gaze contact. Student 2 kept eye contact almost every time she paused.
- Regarding repetition of signs and self-correction, a similar number of repeated signs as well as self-correction were observed in both videos.
- While student 1 produced grammatically correct utterances, utterances of student 2 contained more errors regarding word-order.

The linguistic differences found between both samples were used as explanations of and possible aspects for the perceptions of fluent signing. These aspects lead to seven observational questions about overall fluent signing and its six criteria (1. Pauses 2. Repetitions, 3. Self-corrections, 4. use of meaningless gestures, 5. grammar correctness and 6. text intelligibility.

A group of N = 31 (17 female, 14 male) Deaf native DGS-signers, aged 18 to 60 (mean age: 48) watched the two videos and answered the seven questions by using a 1 to 5 - scale (1 being the best rating).

Results

As expected, Video 1 was rated more fluent than video 2, but surprisingly, the difference between both ratings was very small: Video 1 = *M* 3,29; video 2 = *M* 3,61.

Correlation

Video 1: The perception of fluency correlated with four of the six selected aspects: **highly significant** for intelligibility ($r = 0.59$, $p = 0.001$), self-repairs ($r = 0.58$, $p = 0.001$) and pauses ($r = 0.54$, $p = 0.002$); and **significant** for repetitions ($r = 0.37$, $p = 0.04$).

No other aspect was significantly correlated.

Regarding video 2: **highly significant** for intelligibility ($r = 0.54$, $p = 0.002$) and pauses ($r = 0.47$, $p = 0.007$); and **significant** for repetitions ($r = 0.37$, $p = 0.040$)

No other aspect was significantly correlated, but grammar correctness showed a tendency to significance ($r = 0.33$, $0 = 0.073$).

Principal component analysis

Video1: all six aspects are uploading to a factor (.51 to .88) with repetitions as the weakest value (0.51).

Video 2: two factors are extracted (0.58 to 0.82). Here only repetition seems to be related by 0.84 with a second factor.

Regression analysis

Video 1: self-repairs can predict fluency (b = 0.60, p = 0.017). This factor explains a significant proportion of the variance of fluency, R-square = 0.57, p = 0.001. Video 2: text understanding can be identified as a predictor of fluency (b = 0.46, p = 0.036). This aspect could explain a significant proportion of the variance of fluency: R-square = 0.46, p = 0.014.

Reflection

The general perception and rating of fluency by the Deaf Signers coincided with the preliminary perception of fluency by the two Deaf researchers who selected the videos.

However, based on the linguistic analysis, we expected much higher scores for the first sample and a greater difference between the first and the second video.

On the one hand, the perception of fluency in video 1 was clearly coherent with the six aspects rated in the questionary. On the other hand, fluency in video 2 was perceived based on criteria that were not included in the questionary.

A repetition of the survey with an extended version of the questionary is already planned. Thus, our presentation at the conference can be expected to contain more results than the data the resumed here.

[1] Council of Europe. 2001. *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Univ. of Cambridge.

[2] Oviedo, Alejandro; Reiner Griebel & Leonid Klinner. In progress. „*Sie gebärdet halt nicht sauber*. DGS-Wortschatz und Wahrnehmungen von Flüssigkeit in DGS als L2“. Cologne, University of Cologne.

[3] Kaul, Thomas; Alejandro Oviedo, Reiner Griebel, Leonid Klinner, Tim Prüfer & Monika Krumpen. 2016. C/CSL2, The Cologne Corpus of Sign Language as a Second Language. Poster presented at TaLC 12, University of Giessen, held on 21th July 2016.

[4] Grosjean, François & Harlan Lane. 1977. Pauses and Syntax in American Sign Language. *Cognition* 5: 101-117.

[5] Klima, Edward & Ursula Bellugi. 1979. *The Signs of Language.* Cambridge, MA: Harvard University Press.

[6] Lupton, Linda. 1998. Fluency in American Sign Language. *Journal of Deaf Studies and Deaf Education* 3 (4): 320-328.

[7] Kooij, Els van der, Onno Crasborn & Wim Emmerik. 2006. Explaining prosodic body leans in Sign Language of the Netherlands: pragmatics required. *Journal of Pragmatics* 38: 1598-1614.

Word production difficulties in oral discourse by people with aphasia and healthy speakers: a qualitative analysis[3]

Mariya V. KHUDYAKOVA & Mira B. BERGELSON

National Research University Higher School of Economics

1. Introduction

The concept of a language norm is a complex one, especially when it comes to spoken discourse. Wordfinding problems, disfluencies, speech failures and even grammatical errors occur in spontaneous speech and are in fact its distinct markers which help in the organization of discourse. However, when it comes to analyzing discourse by people with various speech pathologies, it is important to make a distinction between healthy and impaired language output. We propose a qualitative analysis of wordfinding difficulty markers in the speech of neurologically healthy Russian speakers and people with different types of aphasia.

Aphasia is language impairment caused by brain damage, resulting in deficits of production and comprehension of language (Akhutina, 2015; Ardila, 2014; Luria & Hutton, 1977). Word-finding and word-retrieval problems can be very prominent in the speech of people with aphasia (PWA).

2. Material

We analyzed 27 narrative transcripts from the Russian clinical pear stories (Russian CliPS) corpus (Khudyakova et al., 2016). The corpus contains retellings of the Pear film (Chafe, 1980) produced by PWA, people with right hemisphere damage and neurologically healthy speakers of Russian language, recorded and annotated in ELAN tool (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006).

For the current research 27 narratives from the Russian CliPS corpus were chosen: 17 texts by PWA (10 females, mean age – 52.8), and 10 texts by healthy speakers (7 females, mean age – 63.7).

3. Analysis

We performed an analysis of verbal markers of word-finding difficulties. Though pauses (both filled and silent) are annotated in the corpus, they were not taken in the account as their presence, absolute and relative length cannot be attributed to word-production difficulties only. Self-corrections (*apples, no, pears*), false starts (*p= pears*) and discourse markers (*what's the word; let's call it*) were analyzed as markers of word-finding difficulties typical of spontaneous spoken discourse. Only difficulties of noun production were included in the analysis, because verb production difficulties cannot be clearly distinguished from problems with planning of an utterance content or structure (Kibrik & Podlesskaya, 2009, pp. 187–188).

The cases of verbally expressed word-finding difficulties were classified into two large groups: difficulties with retrieval of a phonological word form, and lexical-semantic

3 The abstract was prepared within the framework of the Academic Fund Program at the National Research University Higher School of Economics (HSE) in 2016 (grant №16-05-0024) and supported within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

difficulties. Though errors in semantics and word retrieval are related to different processes, it is not always possible to distinguish the two in a discourse (Howard & Gatehouse, 2006), so we analyze them as one type of lexical-semantic difficulties, which can also reflect discourse planning strategies.

4. Qualitative analysis

4.1 Difficulties on sound form level

Markers of word-production difficulties on the sound form level were observed only in narratives by PWA, with false-starts and self-corrections as markers of word finding. In (1) both markers are present, as well as a common interaction marker *kak_yego* (*what's the word*), which is a verbal expression of inability to find a correct lexeme.

(1) *там тв= твари= как_его товарища угостил*
tam tv= tvari= kak_yego tovarischa ugostil
there cm= cmra= what's_the_word comrade gave_him_a_treat
(comrade) (comrade)

4.2 Difficulties on lexical-semantic and discourse-pragmatic levels

Both in the speech of healthy speakers and PWA we observe word-finding difficulties related to choice between lexical items that are both appropriate in the context. This choice can be attributed to pragmatic level, as in (2), where the healthy speaker is choosing an appropriate word for one of the characters of the story. No semantic error would occur whatever lexeme (*owner* or *gatherer*) is chosen. That's why we call it discourse-pragmatic level difficulties: the problem with the choice of the word is prompted by inadequate knowledge of the world (=pragmatic) or insufficiency of the information in the discourse (=film) itself.

(2) *хозяин за= этот зс= сборщик з= забрался высоко хозяин вот тоже*
hozyain za= etot zs= sborshik z= zabralsya vysoko hozyain vot tozhe
owner cli= this g= gathere cli= climbed high owner well also
(climbed) (gatherer) (climbed)

In examples (3) (healthy speaker) and (4) (speaker with aphasia) we see cases of choice between more and less specific nominations for referents, that can be explained by error on a semantic level or planning on a discourse-pragmatic level.

(3) *но мальчи= дети все едят no malchi= deti vsyo edyat*
but bo= (boys) children everything eat

(4) *плоды это груши plody eto grushi fruit this pears*

Obvious naming errors (choosing a wrong word from the same semantic field) can be observed only in speech by PWA, for example *vegetables* and *apples* as nominations for pears, see (5) and (6).

(5) *пасстанчик эта уносит овощи ой фрукты patsanchik eto unosit ovoshi oj frukty*

guy this carries_out vegetables fruit
oops

(6) *яблоки это нет это как_его*
yabloki eto net eto kak_yego
apples this no this
what's_the_word

5. Preliminary results

As we expected, the majority of speech failures and errors on sound form and semantic level were observed in speech by PWA. The healthy speakers demonstrate word-finding difficulties related to a discourse-pragmatic level, which can possibly be explained by the nature of the task – film retelling. However, it is not always possible to distinguish between disfluencies due to semantic deficits or discourse planning. Though verbal expression of word-finding difficulties in some cases can help to interpret the nature of disfluencies, none of them can be used for unambiguous distinction among phonological, lexico-semantic or pragmatic levels. To distinguish between semantic and discoursepragmatic errors of the healthy participants we will need additional procedures (like debriefing after retellings). As for distinguishing the kind of paraphasias that PWA make (sound form or semantic), it would demand much more detailed diagnosis and case study approach.

Akhutina, T. (2015). Luria's classification of aphasias and its theoretical basis. *Aphasiology*, (August), 1– 20.

Ardila, A. (2014). Aphasia Handbook.

Chafe, W. (1980). *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. (W. Chafe, Ed.). Norwood, New Jersey: Ablex.

Howard, D., & Gatehouse, C. (2006). Distinguishing semantic and lexical word retrieval deficits in people with aphasia. *Aphasiology*, *20*(9), 921–950.

Khudyakova, M. V., Bergelson, M. B., Akinina, Y. S., Iskra, E. V., Toldova, S., & Dragoy, O. V. (2016). Russian CliPS: a Corpus of Narratives by Brain-Damaged Individuals. In *LREC Proceedings*. Portoroz, Slovenia.

Kibrik, A. A., & Podlesskaya, V. I. (Eds.). (2009). *Night Dream Stories: A corpus study of spoken Russian discourse [Rasskazy o snovidenijah: korpusnoe issledovanie ustnogo russkogo diskursa]*. Moscow: Languages of Slavonic Culture.

Luria, A. R., & Hutton, J. T. (1977). A modern assessment of the basic forms of aphasia. *Brain and Language*, *4*(2), 129–151.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a Professional Framework for Multimodality Research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*. Retrieved from https://tla.mpi.nl/tools/tlatools/elan/

Disfluency or speech management? A case study of co-constructed fluency with two Japanese learners of English

Steven KIRK

The University of Tokyo

Although common-sense definitions of fluency, such as those one might find in a dictionary, describe fluent speech with words such as smoothness and flow, actual spontaneous speech is "a highly fragmented and discontinuous activity", where "the attributes of flow and fluency must be judged an illusion" (Goldman-Eisler, 1968, p. 28). Spontaneous speech, as in everyday conversations, contains many examples of what have been traditionally termed *disfluencies*. These include silent and filled pauses, restarts, reformulation, false starts, and repetition. However, Rühlemann (2007), among others, has criticized this characterization of these phenomena, arguing that they are a natural part of spoken language, and only seen as problematic when looked at with the bias toward written language that has been common in linguistic research. Rühlemann also notes that there is no such thing as perfect fluency or "eufluency", with spontaneous speech free of disfluencies, and therefore suggests the more positive term, *speech management*, to describe the words, phrases, and techniques that speakers employ to deal with spoken language in real time.

On the other hand, conversations, co-constructed by two or more speakers, also show signs of smoothness, or what some have called *confluence* (McCarthy, 2009). This co-constructed fluency is achieved by efficient turn exchanges, resulting in the minimizing of gaps between speakers and the general avoidance of overlap—features of conversation that appear to be universal (Stivers et al., 2009). It has been suggested that speakers are able to achieve confluence through alignment at many levels.

For example, speakers align to each other in terms of vocabulary and syntactic use (Garrod & Pickering, 2004), in rhythm of speech (Auer, Couper-Kuhlen, & Müller, 1999; Fiksdal, 1990), and in terms of gaze (Atkinson, 2011; Goodwin & Goodwin, 1986).

The present study aimed to examine the co-construction of fluency in conversation through a parallel case study of two female, university-age, Japanese learners of English. One learner was of lowerintermediate proficiency, and the other was of more advanced proficiency, but both were able to create good impressions of fluent performance on the interlocutors, in conversations with native speakers of English. With a case study methodology, it is possible to conduct an exploratory study that takes into account all of the relevant factors. Non-native speakers were chosen as subjects in order to better facilitate the examination of problems in the co-construction of the conversation. The higher-proficiency learner serves as the baseline to which the lower-proficiency learner is compared, in order to control to some extent extraneous variables.

The case study design was semi-experimental in that it involved a story-retelling task done in monologue and dialogue, which was repeated to take into account the effect of practice. The case study allowed the close examination of the construction of fluency in the story-retelling task moment-by-moment through the course of the retellings, taking into account all relevant factors. The semi-experimental, parallel case study design allowed the findings to be compared (1) between monologue (where the learner recorded herself telling the story alone) and dialogue (where the learner told the story to a native speaker interlocutor), and (2) between the two

learners of differing proficiency. This study was also mixedmethods in that it combined a qualitative approach to data analysis informed by Conversation Analysis, with quantitative comparisons of temporal variables of fluency. It was also multi-modal in that video was employed to take into account gaze, gesture, and head nods.

Results of quantitative analyses revealed that speech of both of the two learners were comparatively more fluent in the dialogues than the monologues in terms of speech rate, articulation rate, and length of silences. However, the higher-proficiency learner had faster speech and articulation rates, and longer uninterrupted runs of speech than the lower-proficiency learner. This implies that narrative in dialogue is not just a listener occasionally backchanneling while the speaker delivers a monologue. The qualitative analyses revealed that the co-construction of smooth conversation was facilitated by the alignment of rhythm between the speaker and listener, supported by gaze, gestures, and head nods.

In terms of speech management, results showed that both learners were able to vary their speech in several ways to deal with difficulties in the retelling of the story. During difficult segments of the retelling, the learners increased the frequency of stressed syllables in relation to the number of words, which allowed them to speak at a lower speech rate while maintaining alignment with the previously established rhythm. For the higher-proficiency learner this was accomplished by shifting from a more nativelike stress-timing to stressing almost every word. For the lower-proficiency learner, who normally stressed almost every word, this was accomplished by stressing multiple syllables in some words, facilitated by adding vowels to the ends of words. In this way, the lower-proficiency learner shifted to more "Japanese" pronunciation of English words in order to slow her speech while maintaining a given tempo of stressed words. This could be interpreted as the learner simply losing control over her pronunciation at points of difficulty, however, at these same points of difficulty in the monologue versions of the story retellings, she did not use this technique, opting instead for much longer silences. This technique of appending vowels to word-final consonants, generally considered to be a pronunciation weakness, has also been shown to be used as a hesitation device by novice Japanese speakers of English (Carroll, 2005). In the present study, this appears to be a choice made by the learner to cope with the difficulties of producing language in real time, and may show that rhythmic alignment is considered a priority for these learners.

These results support previous research that some apparent disfluencies in second language speech should be considered as speech management phenomena, that positively contribute to the coconstruction of fluent conversation. They also suggest that alignment between the speakers in terms of rhythm of speech and gaze may be an important, although perhaps not conscious, goal of conversation. This study has implications for how we view less fluent segments of second language speech, and particularly for the writing of rubrics for evaluating second language spoken fluency.

Atkinson, D. (2011). A sociocognitive approach to second language acquisition: How mind, body, and world work together in learning additional languages. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (1st ed, pp. 143–166). London ; New York: Routledge.

Auer, P., Couper-Kuhlen, E., & Müller, F. (1999). *Language in time: The rhythm and tempo of spoken interaction*. New York: Oxford University Press.

Carroll, D. (2005). Vowel-marking as an interactional resource in Japanese novice ESL conversation. In K. Richards & P. Seedhouse (Eds.), *Applying conversation analysis*. Houndsmills and New York: Palgrave Macmillan.

Fiksdal, S. (1990). *The right time and pace: A microanalysis of cross-cultural gatekeeping interviews*. Norwood, N.J.: Ablex.

Garrod, S., & Pickering, M. (2004). Why is conversation so easy? *Trends in Cognitive Science*, *8*(1), 8–11. Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York: Academic Press.

Goodwin, M. H., & Goodwin, C. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica*, *62*, 51–75.

McCarthy, M. (2009). Rethinking spoken fluency. *Estudios de Lingüística Inglesa Aplicada*, *9*, 11–29. Rühlemann, C. (2007). *Conversation in context: A corpus-driven approach*. London & New York: Continuum.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *PNAS (Proceedings of the National Academy of Sciences)*, *106*(26), 10587–10592.

Should 'Uh' and 'Um' be categorized as markers of disfluency? The use of fillers in a challenging conversational context

Loulou KOSMALA, Aliyah MORGENSTERN

Sorbonne Nouvelle University—Paris 3

'Uh' and 'Um', also known as 'fillers' (Clark & Fox Tree, 2002) are common markers of hesitation in English. They commonly reflect uncertainty (Brennan & Williams, 1995) and lack of confidence, (Fox Tree, 2007) but also serve many other functions such as (1) signaling a delay in speaking (Clark & Fox Tree, 2002); (2) keeping control of the conversational ball (Maclay & Osgood, 1959); (3) marking discourse structure (Swerts, 1998; Tottie, 2014) among many others. They are also mainly associated with notions such as 'difficulty', 'problems' and 'trouble' (Clark, 2006; Finlayson & Corley, 2012; Fraundorf & Watson, 2008; Clark and Fox Tree, 2002) whether they are lexical-based or planning based. Others, however, have argued that 'Uh' and 'Um' have nothing to do with trouble, but rather 'project further talk' (Schlegoff, 2010:140), or serve planning functions (Tottie, 2014) and are thus pragmatic markers.

Speech disfluency is also mainly associated with such values. Speech disfluency, "when fluency breaks down" (Lickley, 2015:12) is a suspension occurring in speech. Such suspension occurs because speakers detect 'trouble' in speech ('Main Interruption Rule', Noteboom, 1983), which makes their speech disfluent; they hesitate, interrupt themselves, repeat phrases, insert a long pause, etc. Fillers such as 'Uh' and 'Um' are also said to be basic forms of disfluency (Shriberg, 1994; Johnson, 1961 quoted by Lickley, 2015).

In order to investigate the function of fillers and whether they are linked to cognitive processing or serve mostly pragmatic functions, we analyzed the use of 'Uh' and 'Um' in the spontaneous speech of English native speakers. We focused on the notion of "challenging contexts" in order to discuss whether fillers should be labeled as markers of disfluency.

We conducted a small experiment involving 16 native speakers of English (8 male and 8 female), aged 18 to 23. The experiment was based on a film, entitled *Big Fish* (2003), directed by Tim Burton. The starting point of our study was to analyze whether fillers were associated

with conceptual difficulty; the film chosen describes the relationship between a father and a son throughout strange and far-fetched stories, and it constantly shifts between fantasy elements and real life events. It is a challenging film, since the viewers always have to make sense of different stories. None of the participants were film specialists or doing film studies.

Eight of the sixteen participants were asked to watch the film a few days prior to the experiment; the experiment was carried out in a series of 8 sessions in which the participants worked in pairs. All the paired participants knew each other fairly well. During each session, the participants who had previously seen the film (coded as Participant A1>A8) answered a list of 10 questions about the film. The questions were read by the other participants (coded as Participant B1>B8) who had not yet seen the film. They were told to be as spontaneous as possible, and to consider the experiment as a casual conversation.

We tried to cover different aspects of the film through the questions; our ultimate goal was to make the questions 'difficult' to answer.

The most difficult questions according to our judgment required that the participants think about concepts such as reality versus fiction and unfamiliar aspects of the film, explain an ambiguous scene and relate their own experiences to what they saw. In order to verify this, we sent a questionnaire to all the participants who had previously answered the questions; they were asked to evaluate the difficulty of each question and rate them from 'very easy to answer' to 'very difficult to answer'.

Each session was video recorded and transcribed with CLAN. The recordings were made in familiar settings (in their homes, or in student halls); the participants were free to speak as much or as little as they wished. The total duration of the corpus was 71 minutes and 63 seconds, with an average duration of 7 minutes and 9 seconds.

We found a total of 347 occurrences of 'uh' (84) and 'um' (223) in the corpus. Despite the more frequent use of 'um', there were no significant differences between the two forms.

Our first striking result was that the answers to the question deemed easiest to answer contained the highest rate of fillers. The answers to the 'difficult' questions, did not contain a significant rate of fillers. This was not consistent with the hypothesis according to which the use of fillers reflected conceptual difficulty. The second major finding was that fillers were more likely to occur near the beginning of utterances rather than in the middle, and to perform the function 'give more time to plan the utterance' (consistent with Tottie, 2014). We also noted that the two answers that received the highest rate of fillers also contained the highest word counts. Therefore, fillers were found to be more often associated with sentence planning, rather than with speech production difficulty.

This led us to two main conclusions: (1) fillers should not necessarily be labeled as markers of disfluency. We argued that fillers do not affect speech the way other markers of disfluency do; while repetitions, selfrepairs, and restarts affect syntax (by breaking one utterance into two, adding or replacing constituents), fillers are added in speech. Since they were found to occur more frequently near the beginning of utterances, they did not necessarily interrupt the flow of speech, but rather gave more time for the speakers to plan the utterance. The second conclusion was that (2) the terms 'difficulty' 'trouble' and 'problems' are problematic, and did not correlate with the use of fillers in our data. We suggested using a less restrictive term— 'effort in discourse planning'—instead (also used by Bortfeld et al. 2001). Our argument is that recalling

from memory, selecting elements from the film, planning the utterance, and conceptualizing an answer are part of a complex process, which involves effort. Therefore, we could argue that fillers occur near the beginning of utterances to facilitate the planning process, in order to formulate the right message.

Bortfeld H; Leon S.D; Bloom J.E; Schober M.F; Brennan S.E (2001) *Disfluency rates in spontaneous speech: Effects of age, relationship, topic, role, and gender*. Language and Speech, Vol 44. Pp 123–147

Brennan, S. E; Williams, M. (1995) *The feeling of another's knowing: prosody and filled pauses as cues to listeners about the metacognitive states of speakers.* Journal of Memory and Language, Vol 34. Pp 383– 398.

Clark, H.H; Fox Tree, J.E. (2002) *Using uh and um in spontaneous speaking*. Cognition, Vol 84. Pp-73-11.

Clark,H.H. (2006) *Pauses and Hesitations: Psycholinguistic Approach*. Encyclopedia of Language and Linguistics, 14-Volume Set, pp.244-248

Finlayson, I.R; Corley, M. (2012) *Disfluency in dialogue: An intentional signal from the speaker?* Psychonomic bulletin & review, Vol 19(5). Pp- 921-928

Fox Tree, J.E. (2007) *Folk notions of um and uh, you know, and like*. Text & Talk, Vol 27(3). Pp-297-314

Frandord, S.H; Watson, D.G. (2011) *The disfluent discourse: Effects of filled pauses on recall*. Journal of Memory and Language, Vol 65. Pp- 161-175

Johnson,W. (1961) *Measurements of oral reading and speaking rate and disfluency of adult male and female stutterers and nonstutterers.* The Journal of Speech and Hearing Disorders, Vol 7. pp 1-20.

Lickley, R. J. (2015). *Fluency and disfluency.* In M. Redford (Ed.), The handbook of speech production (pp. 445-469) Wiley-Blackwell

Maclay, H; Osgood, C.E. (1959). *Hesitation Phenomena in Spontaneous English Speech*. Word, Vol 15(1) pp-19-44

Nooteboom, S. (1980) *Speaking and unspeaking: detection and correction of phonological and lexical errors in spontaneous speech*. In V.A. Fromkin (Ed.), Errors in linguistic performance. New York: Academic Press.

Shriberg, E.E. (1994) *Preliminaries to a theory of speech disfluencies*. Unpublished PhD. thesis, University of California at Berkeley.

Swerts, M. (1998) *Filled pauses as markers of discourse structure*. Journal of Pragmatics, Vol 30. Pp- 485496

Tottie, G. (2014) *On the use of uh and um in American English*. Functions of Language, Vol 21(1). Pp 6-29

Viewing pause behaviour across languages: Methodological and theoretical concerns

Hege LARSSON AAS & Susan NACEY

Inland Norway University of Applied Sciences

This paper discusses methodological and theoretical concerns related to the identification and analysis of unfilled pauses in studies of interlanguage fluency based on material from spoken learner corpora. It addresses the following research question:

How can a spoken learner corpus be compiled and used to make valid claims about pause behaviour, and in what ways can native language material inform contrastive analyses of inter- and target language speech?

The discussion posits as its point of departure a description of the process of compiling a spoken corpus of learner interviews, the hitherto unpublished Norwegian version of the Louvain International Database of Spoken English Interlanguage (LINDSEI) (Gilquin, De Cock, & Granger, 2010), as well as its smaller counterpart consisting of interviews with some of the same speakers speaking in their L1 Norwegian (*NL1,* cf. Gilquin (2008)). It addresses challenges involved at all stages of collecting and preparing the interlanguage and NL1 material, and considers what research can be conducted with such spoken learner data in order to generate valid results, with a particular emphasis on pause behaviour across languages. A survey is presented of the procedure for segmentation of speech into turns and utterances in the original LINDSEI transcriptions, and its possible implications for the annotation and analysis of pauses. These conventions are contrasted with a new suggestion for transcription and annotation practices that addresses some of the issues presented.

The main purpose of the paper is thus twofold: 1) it highlights larger issues that are considered crucial for the validity of learner corpus data in general, such as the inclusion of same-speaker NL1 material and considering the cross-linguistic perspective, and 2) it suggests transcription and annotation practices that may contribute to the validity of future research on pause behaviour using large-scale corpus data.

The LINDSEI corpus consists of corpora of interviews with non-native speakers of English from different native language backgrounds, and these corpora have previously been investigated in comprehensive studies of fluency features in non-native speech (e.g. Brand & Götz, 2011; Götz, 2013), often in contrast with the comparable target language (*NL2*) corpus LOCNEC (De Cock, 2004). From the perspective of interlanguage fluency and disfluency research, the relative presence of unfilled pauses is often viewed as primarily a reflection of speakers' processing constraints, and between-group comparisons with NL2 behaviour could be seen to highlight this view of interlanguage pausing. Looking at pause behaviour from the perspective of fluent native language production, on the other hand, the presence of pauses in a person's speech is more likely to be seen as a reflection of personal style (Fillmore, 1979), and native speech is often "perceived as fluent despite all the pausing that normally comes with it" (Riazantseva, 2001, p. 500). However, the fact that pauses can serve a multitude of functions unrelated to the stalling of time during speech management in native speech should also open up for a view of interlanguage pause behaviour as a reflection of "textual idiosyncrasies characteristic of individual speaking style" (Raupach, 1980, p. 264). Although corpus-based studies of transfer have traditionally focused on the possible transfer of lexical or syntactic patterns, some studies indicate that interlanguage pause behaviour may also be closely related to native language or culture (e.g. Lehtonen, 1979, Riazantseva, 2001). This presents the possibility that differences found between interlanguage and target language speech in these areas may be partly explained in terms of transfer, or "the influence that previous knowledge or skills have on future learning" (Osborne, 2015, p. 333). In light of these and similar concerns, this paper argues that the potential for transfer should be a central component in the interpretation of pause behaviour in studies of interlanguage speech, and that the collection of comparable NL1 data should form an integral part of the compilation of interlanguage speech corpora. The main issues involved in ensuring comparability in the compilation of cross-language interview material are also discussed.

Our view of pause behaviour across languages is further constrained by our choices at the transcription stage of the corpus compilation. In the LINDSEI transcription system any

utterance is treated as an interruption of the current turn holder's turn (Gilquin et al. 2010), resulting in the interruption of speech in sequences like the following (text enclosed by <B></B> belong to the informant, and text enclosed by <A></A> belong to the interviewer):

1. <B> (eh) it's a: . nice town . (eh) it's (eh) it's a very touristy . I find </B>

<A> (mhm) </A>

<B> it's lots and . because we were there when there was a holiday too so (eh) it was very crowded . everywhere </B>.

Here, the silence at the end of B's first utterance is not transcribed, and hence would not be included in an analysis of this speaker's pause behaviour. As observed by Du Bois et al. (1992), "in some cases, the question of who a pause belongs to, how long it lasts, and even whether it has occurred in a specific place, become subtly and inextricably linked to the interpretation of turn-taking and overlapping between speakers" (p. 42). In an attempt to bring this perspective to the forefront, alternative transcription conventions are suggested, involving the segmentation into turns and utterances according to a set of criteria which includes discriminating between contributing and non-contributing utterances (cf. Linell & Gustavsson, 1987).

This paper thus argues that just as speech production as a whole cannot be considered in isolation, utterances should not be viewed as independent from their immediate co-text. The segmentation approach presented here is a step towards combining a dialogical analysis with an exploration into a specific fluency variable, which in turn may contribute to a more comprehensive view of fluency in both native- and interlanguage speech.

Brand, C., & Götz, S. (2011). Fluency versus accuracy in advanced spoken learner language: A multimethod approach. *International Journal of Corpus Linguistics, 16*(2), 255-275.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures (BELL), New Series, 2*, 225-246.

Du Bois, J. W., Cumming, S., Schuetze-Coburn, S., & Paolino, D. (Eds.). (1992). *Santa Barbara Papers in Linguistics, Discourse transcription* (Vol. 4).

Fillmore, C. J. (1979). On Fluency. In C. J. Fillmore, D. Kempler, & W. S.-Y. Wang (Eds.), *Individual Differences in Language Ability and Language Behaviour* (pp. 85-101). New York: Academic Press.

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.

Götz, S. (2013). *Fluency in Native and Nonnative English Speech*. Amsterdam: John Benjamins.

Lehtonen, J. (1979). Speech rate and pauses in the English of Finns, Swedish-speaking Finns, and Swedes. In R. Palmberg (Ed.), *Perception and production of English: Papers on Interlanguage.* (pp. 35-51). Åbo: Åbo Akademi.

Linell, P., & Gustavsson, L. (1987). *Initiativ och respons: Om dialogens dynamik, dominans och koherens*. Linköping: Department of Communication Studies.

Osborne, J. (2015). Transfer and learner corpus research. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 333-356). Cambridge: Cambridge University Press.

Raupach, M. (1980). Temporal variables in first and second language speech production. In H. W. Dechert & M. Raupach (Eds.), *Temporal Variables in Speech. Studies in Honour of Frieda Goldman-Eisler* (pp. 263-270). The Hague: Mouton Publishers.

Riazantseva, A. (2001). Second language proficiency and pausing. *Studies in Second Language Acquisition, 23*, 497-526.

Stuttered and non-stuttered disfluencies in normally fluent, French-speaking preschool children

A.-L. LECLERCQ & A. MOYSE

University of Liège

The criterion of 3% stuttering-like disfluencies (SLD) – (part-)word repetitions, sound prolongations and blocks (i.e. inappropriate stoppages of the flow of air or voice) – is often suggested for diagnosing the stuttering in children (e.g., Boey et al., 2007; Yairi & Ambrose, 2005). There is a longlasting debate among scholars about the relevance of considering monosyllabic word repetitions as stuttering-like disfluencies, given their high frequency in typically developing children (e.g., Howell, 2013; Wingate, 2001). Yet, these are prime characteristics that prompt identification of early stuttering by parents (Yairi & Ambrose, 2013). Nevertheless, there is very little normative data concerning the disfluencies occurring in the speech of normally fluent children (Tumanova et al., 2014), and there is no data on French speakers. The aim of this study is to examine the need for reference data in French in order to distinguish typical developmental disfluency from stuttering.

We used the CHAT transcription system and the coding conventions of FluencyBank to transcribe and analyse the speech disfluencies of twenty monolingual, French speaking children who do not stutter, aged 4 years. FluencyBank (Bernstein-Ratner & MacWhinney) is a shared database for the study of the development of fluency in both normal and disordered populations. It is a project of Talkbank that seeks to archive and facilitate sharing of data relevant to fluency research. In the global word count, we only took into account the intended words (i.e. not the interjections) and the final formulation/sentence structure for the revised sentences. We wanted to reflect how the final message of the child was disfluent – in other words, how much disfluencies are necessary before the child arrives at his core message.

Results reveal higher percentages of disfluencies than previously observed for English**,** with an average of 10% total disfluencies, and around 8% non-stuttered disfluencies. As expected, SLD occur less frequently than 3 in 100 words (around 2%), but there is a high variability among children (up to 7%), mainly due to the frequency of monosyllabic word repetitions. However, most are repeated fewer than three times. Our results support the need to take caution when considering monosyllabic word repetitions as stuttering-like disfluencies.

Boey, R., Wuyts, F., Van de Heyning, P., De Bodt, M., & Heylen, L. (2007). Characteristics of stutteringlike disfluencies in Dutch-speaking children. *Journal of Fluency Disorders, 32*, 310-329.

Howell, P. (2013). Screening school-aged children for risk of stuttering. *Journal of Fluency Disorders, 38*, 102-123.

Tumanova, V., Conture, E., Lambert, E., & Walden, T. (2014). Speech disfluencies of preschool-age children who do and do not stutter. *Journal of Communication Disorders, 49,* 25-41.

Wingate, M. (2001). SLD is not stuttering. *Journal of Speech, Language and Hearing Research, 44,* 381-383.

Yairi, E., & Ambrose, N. (2005). *Early childhood stuttering.* Austin, TX: Pro-Ed.

Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st advances. *Journal of Fluency Disorders, 38,* 66-87.

(Dis)fluency in EFL read-aloud: The orthographic context

Signe-Anita LINDGRÉN

Åbo Akademi University, Aalto University

The present oral paper shows statistically significant differences between a group of dyslexic and non-dyslexic university students in fluency marker frequency and types, the orthographic context of the disfluencies, and the positions in a sentence in an EFL (English as a foreign language) reading-aloud task and discusses the pedagogical implications and the relevance of the findings for (corpus) research on EFL (dis)fluency.

The study explores in particular the placement of hesitations and pausing (Warren, 2012; Watanabe & Rose, 2012) in the oral performance of a specific group of EFL users and learners, namely those whose performances are affected by dyslexia. Dyslexia, or specific reading and writing disabilities, (e.g. Høien & Lundberg, 2000; Lyon et al, 2003; Vellutino et al, 2004) is considered the most common handicap in the Western world (von Euler, 1996) and affects 5-15% of the population (e.g. Brunswick, 2010), and, even though reading in dyslexics can improve over time, specific features related primarily to decoding and fluency keep distinguishing reading performances between dyslexics and non-dyslexics (Høien & Lundberg, 2000). Further, it is well established that dyslexia affects the performances in foreign languages (e.g. Ganschow & Sparks, 2001), observed also in highly advanced language users (e.g. Lindgrén, 2012; Lindgrén & Laine, 2011). However, how it surfaces in FL learners of different L1s is still an under-researched, yet, essential question (cf. Ziegler & Goswami, 2005).

This paper analyses in further detail an EFL reading-aloud task (see e.g. Wengelin 2002 on a corresponding written task on a different group of dyslexic participants) collected as part of a five-hour test battery administered to twenty university students with dyslexia (L1 Finland-Swedish) and a chronological age-matched and education-matched control group (n=20) with the same mother tongue background (Lindgrén, 2012). The test battery contained various reading and writing tasks in Swedish, Finnish, and English, as well as further cognitive tasks relevant for dyslexia (Lindgrén, 2012; Lindgrén & Laine, 2011). In the EFL read-aloud, the participants were asked to read three unfamiliar passages with a correct, normal pronunciation and intonation, preferably at a good speed, yet at their own pace. The performances were recorded and analyzed offline. Both quantitative and qualitative analyses were performed; partly using PRAAT, observing also, for instance, phrase boundaries. The statistical analyses applied non-parametric tests (e.g., Mann Whitney U-Test, the Wilcoxon test).

A previous study investigating errors, hesitations and pauses in these data showed that dyslexia may be reflected in deviating error types and error frequencies, pause types and placements with statistically significant group differences (Lindgrén & Laine, 2011; Lindgrén & Mattsson, 2013). The question arises whether the differences in the disfluency markers (types and frequency) are related to orthographic characteristics of linguistic entities previously shown to be difficult for dyslexics, e.g. homophones and morphophonological features (Lindgrén & Laine, 2011; Moats, 1993, 1996; cf. Carney, 1994; van Berkel, 2005). The placement, and thus the orthographic context of the disfluency phenomena, is therefore in focus in the proposed paper. As for the task type itself, using a reading-aloud is in line with Gibson (2008) and Nation (e.g. 2009), who advocate for the (re)introduction of reading-aloud tasks in the FL classroom and point to several learning benefits of this type of activity (cf. Ziegler & Goswami, 2005). Exploring how advanced dyslexic EFL readers fare in detail in such a task will cast light on whether dyslexic EFL learners would benefit from targeted decoding instruction, and especially which linguistic features need to be attended to (cf. Schneider, 2016). The results may also have implications for the interpretation of results based on corpus studies in a more general scope.

In the oral paper, I first briefly present the theoretical framework and selected previous findings, the participants of the study, the data collection and the task. I then concentrate on the analyses and the results. The findings will be discussed in the light of pedagogical implications and relevance for (corpus) research on EFL (dis)fluency.

The results of the present study advance our understanding of the EFL learner and of dyslexia and foreign language learning.

Berkel van, A. 2004. Learning to spell in English as a second language. *IRAL*, 42: 239–257.

Berkel van, A. 2005. The role of the phonological strategy in learning to spell in English as a second language. In V. J. Cook (Ed.), *Second language writing systems* (pp. 97–121). Clevedon: Multilingual Matters.

Brunswick, N. 2010. Unimpaired reading development and dyslexia across different languages. In N. Brunswick, S. McDougall, & P. de Mornay Davies (Eds.), *Reading and Dyslexia in Different Orthographies* (pp. 131–154). Hove: Psychology Press.

Carney, E. 1994. *A survey of English spelling*. London: Routledge.

Gibson, S. 2008. Reading aloud: a useful learning tool? *ELT Journal, 62:* 29–36.

Götz, S. 2013. *Fluency in Native and Nonnative English Speech.* Amsterdam; Philadelphia: John Benjamin Publishing Company.

Ganschow, L. & Sparks, R. L. (2001). Learning difficulties and foreign language learning: A review of research and instruction (review article). *Language Teaching 34*: 78–98.

Høien, T. & Lundberg, I. (2000). *Dyslexia: From Theory to Practice.* Dordrecht NL, Kluwer. Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia 53*, 1–14.

Lindgrén, S-A. 2014. What can we learn from linguistic features like pauses and hesitations in reading aloud tasks?" Paper presentation. The British Dyslexia Association International Conference, BDA, Guildford, UK.

Lindgrén, S-A. 2012. *Mild Developmental Dyslexia in University Students: Diagnosis and Performance Features in L1, L2, and L3*. English Department Publications 6. Turku, Åbo Akademi University.

Lindgrén, S-A. & Laine, M. 2011. Multilingual dyslexia in university students: Reading and writing patterns in three languages. *Clinical Linguistics & Phonetics, 25,* 753–766.

Lindgrén, S-A. & Mattsson, D. 2014. Hesitations in EFL reading among advanced students with and without dyslexia." Interactive poster. Autumn Symposium of the Finnish Association of Applied Linguistics, AFinLA. Turku/Åbo.

Moats, L. C. 1993. Spelling error interpretation: Beyond the phonetic/dysphonetic dichotomy. *Annals of Dyslexia,43*, 174–85.

Moats, L. C. 1996. Phonological spelling errors in the writing of dyslexic adolescents. Reading and Writing: *An Interdisciplinary Journal, 8*, 105–119.

Nation, P. 2009. Reading faster. *International Journal of English Studies 9*(2): 131–144.

Schneider, E. 2016. Let's make sense of the English spelling mess to help all learners including dyslexics. Paper in session Lindgrén, S-A. How to support dyslexics in foreign language learning. 3rd Baltic Sea – 17th Nordic Literacy Conference, Åbo/Turku.

Schneider, E. (1999). *Multisensory Structured Metacognitive Instruction: An Approach to Teaching a Foreign Language to At-Risk Students*. Frankfurt am Main: Peter Lang.

Vellutino, F. R., Fletcher J. M., Snowling, M. J. & Scanlon, D. M. 2004. "Specific reading disability (dyslexia): What have we learned in the past four decades?" In *Journal of Child Psychology and Psychiatry* 45 (1): 2–40.

Warren, P. 2013. *Introducing Psycholinguistics.* Cambridge: Cambridge University Press.

Watanabe, M. & Rose, R. L. 2012. Pausology and hesitation phenomena in second language acquisition. In Robinson, P (ed.). *The Routledge Encyclopaedia of Second Language Acquisition.* New York/London: Routledge: 480-483. Pre- publication version.

Wengelin, Å. 2002. *Text Production in Adults with Reading and Writing Difficulties.* Göteborg: Department of Linguistics.

Ziegler, J. C. & Goswami, U. 2005. Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory". In *Psychological Bulletin* 131: 3–29.

Exploring how to support novice academic EFL presenters: A case study on teddy bears and hesitations

Signe-Anita LINDGRÉN[1,2] & Kaisa KORPINEN[1]

Åbo Akademi University[1], Aalto University[2]

High-stake presentations in front of class raise the affective filter especially in novice presenters and in a foreign language setting and can have an unwanted influence on the quality of the language produced, and, in end effect, on the message transmitted, or rather how it is interpreted by the hearer. Presentation skills are general skills that individuals with an academic degree are often taken for granted to master. For this reason, among others, they are part of the curriculum and also deserve scientific scrutiny and awareness-raising in the classroom.

The present case study explores the use of formulaic language and the type and frequency of lexical teddy bears, repeats, self-corrections, and filled pauses in prepared, but free speech. The aim of the study is to gain insights into novice academic students' performances in order to adapt the EFL teaching and instructions regards oral presentations based on research-

informed knowledge specific to the proficiency group, learning context, culture and language background (L1 Swedish), thus conforming to action research.

The data used were collected (video and audio recorded) at the end of the first year of university studies in English in a course focusing on English for professional purposes (spring 2016; Lindgrén, 2016). In this data set, the students gave a 10-minute persuasive group presentation with clear instructions, for instance, on peer evaluation, which focused on the contents and the effect of the arguments presented (i.e. not related to the features investigated in this study). After the presentation, the participants answered various questions related to previous experiences in giving academic presentations and on attitudes towards academic speech. (Written consent for research and was received from all participants). In the poster, the data from three of four groups, each with four participants, are analyzed qualitatively through close-reading and quantitatively with statistical measures. The oral data have been transcribed in two ways: one conforming to orthographic rules in order to facilitate the use of Wordsmith, the other observing details needed for repairs, selfcorrections and filled pauses. Transcriber, Wordsmith, and SPSS are used.

In the poster, we briefly present the theoretical framework and our research questions, the setting and the task, and focus on the findings of the linguistic analyses and the questionnaire, as well as the discussion of the pedagogical implications. In the poster session, we gladly present also relevant previous research in more detail.

Carey, R. 2013. On the other side: formulaic organizing chunks in spoken and written academic ELF. *Journal of English as a Lingua Franca, 2*(2): 207-228.

Ellis, R. 2012. *The Study of Second Language Acquisition.* 2nd ed. Oxford: OUP.

Gilquin, G. 2015. The use of phrasal verbs by French-speaking EFL learners. A constructional and collostructional corpus-based approach. *Corpus Linguistics and Linguistic Theory*, *11*(1): 51-88.

Götz, S. 2013. *Fluency in Native and Nonnative English Speech.* Amsterdam/Philadelphia: John Benjamin Publishing Company.

Hasselgren, A. 1994. Lexical teddy-bears and advanced learners: a study into the ways

Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics, 4*(2): 237–260.

Hilton, H. 2008. The link between vocabulary knowledge and spoken L2 fluency, *The Language Learning Journal, 36*(2): 153-166.

Lennon. P. 1990. Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*(3): 387-417.

Lindgrén, S-A. 2016. Talkabulary data. (unpublished.) English Language and Literature. Åbo Akademi University. Åbo/Turku.

Lintunen, P. & Skaffari, J. 2014. L2 English learners as public speakers: generic and languagespecific features in learner narratives. *AFinLA-e, 6.*

McCarthy, M. 2010. Spoken fluency revisited. *English Profile Journal, 1*(1): 1-15.

Meunier, F. 2012. Formulaic language and language teaching. *Annual Review of Applied Linguistics, 32:* 111-129.

Paquot, M. & Granger, S. 2012. Formulaic language in learner corpora. *Annual Review of Applied Linguistics, 32:* 130-149.

Peltonen, P. & Lintunen, P. 2015. Integrating quantitative and qualitative approaches in L2 fluency analysis: A study of Finnish-speaking and Swedish-speaking learners of English at two school levels. *European Journal of Applied Linguistics, (4)*2: 209-238.

Rossiter, M. J. 2009. Perceptions of L2 fluency by native and non-native speakers of English. *The Canadian Modern Language Review, 65*(3): 395-412.

Ruhi, S, Haugh, M., Schmidt, T. & Wörner, K. 2014. *Best Practices for Spoken Corpora in Linguistic Research*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Schmitt, N. (ed). 2004. *Formulaic sequences*. Amsterdam: John Benjamins Publishing Company.

Simpson, R. C. 2004. Stylistic features of academic speech: The role of formulaic expressions In Connor, U. & Upton, T.A. (eds.) *Discourse in the Professions: Perspectives from Corpus Linguistics*. (PP. 37-64.) Amsterdam/Philadelphia*:* John Benjamin Publishing Company.

Simpson-Vlach, R. & Ellis, N. 2010. An academic formulas list: New methods in phraseology research. *Applied Linguistics, 31*(4): 487-512.

Making invisible "trouble" visible: Self-repair increases abstraction in dialogue

Gregory MILLS & Gisela REDEKER

University of Groningen

A central finding in dialogue research is that interlocutors rapidly converge on referring expressions which become progressively contracted and abstract. This occurs for a wide range of referents, e.g. when referring to spatial locations (Garrod and Doherty, 1994; Roberts et al., (2016), music (Healey et al., 2007), concepts (Schwartz, 1995), confidence (Fusaroli et al., 2012), and temporal sequences (Mills, 2011). Cumulatively, these findings suggest that interaction in dialogue places important constraints on the semantics of referring expressions.

However, there is currently no consensus about how best to account for how convergence develops. The iterated learning model of Kirby and Hurford (2002) explains convergence as arising out of individual cognitive biases; the interactive alignment model of Pickering and Garrod (2004) favours alignment processes, while the collaborative model of Clark (1996) emphasizes the role of positive feedback. By contrast, Healey et al., (2007) argues that negative evidence of understanding plays the central role: When interlocutors initiate repair, this allows them to interactively identify, diagnose and resolve any differences in interpretation between them and their conversational partner. Addressing these differences accelerates convergence.

To investigate in closer detail how negative evidence contributes toward convergence, we report a variant of the "maze task" (Pickering and Garrod, 2004). Over the course of an hour, pairs of participants attempt to solve twelve mazes. Participants communicate with each other via an experimental chat tool (Healey and Mills, 2006), which automatically transforms participants' private turn-revisions into public self-repairs that are made visible to the other participant. For example, if a participant, A types:

A: Now go to the square on the left, next to the big block on top

and then before sending, A revises the turn to:

A: Now go to the square on the left, next to the third column

The chat server automatically detects the revised text and inserts a hesitation marker (e.g. "umm" or "uhhh" immediately preceding the revision). This would yield the following turn, sent to B:

A: Now go to the square on the left next, to the big block on top umm.. I meant next to the third column

Two self-repair formats were used:

(1) A: original turn + hesitation marker + reformulated turn

(2) A: original turn + hesitation marker + "I meant" + reformulated turn

To avoid triggering interventions for simple typing errors, only turns which deleted more than one word were transformed into public self-repairs. Interventions were performed symmetrically on both members of a dyad. No participants detected the experimental manipulation.

Examining the transcripts showed that participants who received these transformed turns used more abstract Cartesian location descriptions than participants in a baseline condition. This pattern was already apparent after 5 minutes in the task. Task performance followed a different pattern – initially participants who received these interventions performed worse – completing fewer mazes and requiring more moves to solve each maze. However, by the end of the task, participants who received the interventions performed at the same level as participants in the baseline condition. Crucially, participants who received transformed turns continued to use more abstract descriptions.

We argue that this effect is due to the artificial self-repairs having a beneficial effect of amplifying naturally occurring signals of miscommunication (cf. Healey et al, 2007): the artificially generated disfluencies and reformulations are used by participants as cues that their partner is having difficulty coordinating on the semantics of referring expressions. Consequently, participants expend more effort to address these putative problems – and once these problems have been identified and resolved, dyads are able to converge quicker on more stable and more abstract referring schemas.

Clark, H. H. (1996). Using language. 1996. Cambridge University Press: Cambridge/

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to terms quantifying the benefits of linguistic coordination. Psychological science, 0956797612436816.

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. Cognition, 53(3), 181-215.

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, *1*(2), 292-304.

Healey, P. G., Swoboda, N., Umata, I., & King, J. (2007). Graphical language games: Interactional constraints on representational form. Cognitive Science, 31(2), 285-30

Kirby, S., & Hurford, J. R. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In *Simulating the evolution of language* (pp. 121-147). Springer London.

Mills, G. J. (2011, September). The emergence of procedural conventions in dialogue. In Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue. Cognitive Science Society (pp. 210-211).

Roberts, G., Lewandowski, J., & Galantucci, B. (2015). How communication changes when we cannot mime the world: Experimental evidence for the effect of iconicity on combinatoriality. *Cognition, 141,* 52-66.

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. The Journal of the Learning Sciences, 4(3), 321-354.

Vocal ticking in Tourette's syndrome

Mairym LLORÉNS MONTESERÍN

University of Southern California

1. Introduction

Tourette's syndrome (TS) is a neurological condition that affects 1% of children worldwide [1] and in severe cases can persist into adulthood [1-3]. The condition is characterized by the presence of an inventory of abrupt, repetitive movements and vocalizations called tics. *Vocal ticking*—the focus of this paper—is production of tics that are executed through movements of the vocal tract articulators and result in audible sound. Linguists and speech scientists are uniquely equipped to investigate vocal ticking. Clinically motivated research on TS can make use of linguistically-informed analyses of vocal tics to identify mechanisms underlying TS neuropathology. In addition, vocal tics and vocal ticking patterns are of interest to linguists and speech scientists because they speak to the typical functioning of the speech production system, language and dis-fluencies. Vocal tic production engages the larynx and other speech articulators and co-exists with fluent speech, but vocal tics appear unlike speech in the sense that they are incoherent with respect to speech prosody. This point will be elaborated below. In addition, the distribution of vocal tics with respect to certain dis-fluencies will be discussed. Pauses, for instance, are typically characterized by the cessation of articulator movement but in TS speech vocal tics can emerge at locations where pauses are expected and observed to occur. They also occur around other types of dis-fluencies such as false starts and repeats or repairs.

This paper reports preliminary findings from linguistically-informed analyses of tic and non-tic vocal behavior by a single, female, adult English-speaker with TS that were aimed at addressing the following questions: does tic "speech" pattern at all like fluent speech? If not, does it pattern with one kind of dis-fluency, the filled pause? The utterances analyzed were produced by the speaker during various interviews and performances that were recorded and posted online. One notable feature of this particular speaker's TS is the highly frequent production of a small inventory of tic words. For instance, the tic word "biscuit" in one recording was produced an average of 11 times per minute on a background of ongoing and apparently normal voluntary speech. This feature of the speaker's TS motivated a comparative analysis of her intonation during tic and non-tic words as well as an investigation into the discourse structure of her non-interactive utterances in order to identify where her tic words appear with respect to the non-tic words around them. Two notable observations can be reported at this stage. First, scrutiny of the speaker's intonation suggests that tic word melody *interrupts* ongoing sentential melody rather than adhere or scale to it. Such a finding is suggestive of a contrast between vocal tics and filled pauses, whose intonation has been previously found to scale to that of the surrounding sentential context [4]. Second, analysis of the texts of non-

interactive utterances by the same speaker revealed that tic words did not surface during segments of fluent speech but did occur *between* fluent segments (i.e. during pauses), as well as during dis-fluent speech (i.e. during repetitions and repairs). These observations are discussed in more detail below.

2. Tic and non-tic speech behavior in a person with Tourette's

Several Youtube videos of a single female TS patient were identified, downloaded and transcribed. Audio was extracted from the downloaded videos for analysis in Praat [5]. Recordings were segmented into tic and non-tic portions. F0 contours were extracted automatically in Praat. As expected, the speaker's intonation during unambiguously sentential portions is typical for any speaker of her variety of English. F0 during production of tic words, however, departs strikingly from intonation during the sentential elements surrounding the tics. These sudden shifts in the speaker's f0 appear abrupt and out of place.

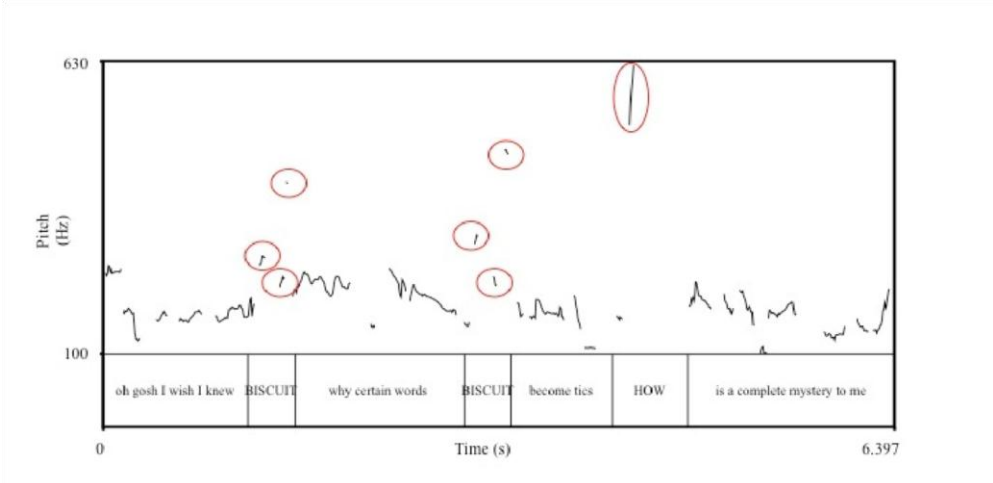


Figure 1. Exemplar of spontaneous utterance by TS speaker with automatically-extracted pitch contours. Red circles surround pitch portions that correspond to three vocal tics produced. Transcription of utterance, with tics in all caps: "Oh gosh I wish I knew BISCUIT why certain words BISCUIT become tics HOW is a complete mystery to me".

If the utterance represented in Figure 1 is played back to a naïve listener with the tic portions cut out, she will not notice that anything is missing. To put another way, the utterance's prosody is coherent with the tics cut out but tic prosody is incoherent with the sentential elements around it. It is in this sense that production of tic words appears unlike speech, despite the segmental similarity of these tics to lexical items such as "biscuit".

Transcriptions of the speaker's utterances in each recording were also subjected to analysis using the linguistic discourse model (LDM) of Livia Polanyi and MartinVanDerBerg.[4] This model generates a single hierarchical tree structure for swaths of text after breaking up the text into segments or phrases according to a combination of the semantic, syntactic and discourse structure of the entire utterance. Consider the following segments of one such tree, with pauses

4 Transcriptions—but not recordings—were shared with Polanyi and VanDerBerg via personal communication for naïve segmentation.

annotated in parentheses and tic words in all caps. The text represents the speaker's response to a question during an interview.

1 they were m-much BISCUIT much less noticeable HOW HOW (pause)

2 when I was younger (pause)

3 than they are now (pause)

4 less noticeable to other people (pause) BISCUIT (pause)

5 the sensation hasn't actually changed BISCUIT (pause)

6 that much for me BISCUIT

Fluent segments of speech in this type of analysis are those text sequences like (4) that do not encompass pauses, repeats or repairs. It is the case that vocal tics can occur *following* fluent segments, that is, at locations where filled and unfilled pauses are also observed to occur in both the speech of neurotypicals and in the voluntary speech of the TS speaker. Indeed, preliminary analysis of audio corresponding to this particular text confirms that vocal tics are flanked by pauses. In contrast, segments like (1) containing false starts and repeats can be thought of as dis-fluent segments. Vocal tics also do occur *during* dis-fluent segments like (1). In other words, vocal tics can be flanked by pauses as well as repeats or repairs.

3. Concluding Remarks

Research on the production of TS tics and on the properties of those tics is in its infancy. Preliminary assessment of available data strongly suggests that in a person with TS, vocal production of voluntary speech and vocal production of tic "speech" constitute two different types of action—regardless of the apparent lexical content that an observer may attribute to tic words. Studies are currently underway to further demonstrate that from the perspective of the listener, an utterance by a person with TS with any vocal tics removed is perceived and evaluated as complete and "normal". Also currently underway is an attempt to quantify the claim that tic word intonation, unlike intonation during filled pauses, cannot be predicted from intonation during the non-tic speech surrounding tics. Such a finding would be an experimental verification of the role of voluntary communicative intent in generating speech prosody and thus contribute to a more nuanced understanding of the genesis of prosodic structure. It would further help classify filled pauses, unlike vocal tics, as part of healthy, fluent speech production, a potentially significant advancement in the field of dis-fluencies. Furthermore, this and other basic research is necessary if the clinical understanding of TS is to move beyond its designation as a movement disorder and toward a mechanistic account which would allow for the development of TS-specific therapeutic targets.[5]

1. Robertson MM, V Eapen and AE Cavanna (2009) International prevalence, epidemiology and clinical phenomenology of Tourette syndrome: a cross-cultural perspective. Journal of Psychosomatic Research 67, pp475-483.

2. Singer, HS (2005) Tourette syndrome: from behavior to biology. Lancet Neurology 4, pp149159.

5 TS is classified as one of the motor tic disorders, themselves classified as neurodevelopmental disorders in the DSM-V

3. D Martino, N Madhusudan, P Zis, AE Cavanna (2013) An introduction to the clinical phenomenology of Tourette syndrome. International Review of Neurobiology 112, pp1-34.

4. EE Shriberg (1994) Preliminaries to a theory of speech disfluencies. Doctoral dissertation, University of California at Berkeley.

5. Boersma, Paul & Weenink, David (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.23, retrieved 12 December 2016 from http://www.praat.org/

6. American Psychiatric Association (2013). *Diagnostic and statisical manual of mental health disorders* (5th ed.). Washington, DC: Author.

Discourse markers within (dis)fluent constructions in English, French and Spanish casual conversations: The challenges of contrastive fluency research

Elena PASCUAL[1], Ludivine CRIBLE[2]

Universitat de València[1], Université catholique de Louvain[2]

Casual conversation is the most natural and spontaneous form of human communication, where speech is subject to the constraints of temporality and immediate communicability. These restrictions mean that discourse planning takes place as speakers go along (Briz 1998). This *on-line* character of spoken language (Auer 2009) is manifested in speech production and comprehension in a number of distinct markers such as repetitions, fresh-starts, unfilled pauses or discourse markers. Götz (2013) characterizes these markers as 'fluencemes', a term that conveys their role in contributing to the fluency or disfluency of discourse. This study focuses on one of these (dis)fluent phenomena, namely discourse markers (henceforth DMs), and examines their co-occurrence with other fluencemes across three different languages in two comparable corpora: *DisFrEn* (Crible under review), a dataset that contains 17,000 words of spoken conversations in English and in French; and a sample of 17,000 words from *Val.Es.Co 2.0* (Cabedo and Pons 2013), a corpus consisting of spoken conversations in Spanish.

We start from the hypothesis that there are recurrent patterns of sequences of disfluencies containing DMs. Our objective is to uncover and contrast these patterns in English, French and Spanish with the overall aim of determining if the DMs present in these sequences are formally and functionally equivalent across the mentioned languages.

To carry out a contrastive analysis of DMs and disfluencies across languages is no easy undertaking. DMs are a complex category that groups together heterogeneous elements, for example, conjunctions (*so*, Fr. *donc*, Sp. *pues*), adverbs (*well*. Fr. *bon*, Sp. *bueno*) and verbal phrases (*you know*, Fr. *tu sais*, Sp. *sabes*). They are subject to a high variability in terms of their syntactic and functional status. With regard to syntax, DMs are highly flexible in terms of position to the extent that they are in many cases optional (Schiffrin 1987). DMs are, moreover, polyfunctional and context-sensitive. The same form can carry out diverse functions depending on the context or even multiple functions simultaneously. These features make DMs complicated linguistic devices that are particularly challenging for cross-linguistic studies, especially when attempting to find common ground among multilingual corpora that employ distinct theoretical frames for annotating the categories and functions of DMs (see, for example,

the varying approaches pursued in the Penn Discourse Tree Bank 2.0 [Prasad et al. 2008] and in Rhetorical Structure Theory [Mann and Thompson 1988]). The lack of consensus and use of divergent annotation conventions also hinder the contrastive study of disfluencies (cf. Shriberg 1994 and Pallaud et al. 2013, among other annotation proposals).

This study aims to tackle these complexities – the use of dissimilar annotation schemes for disfluencies in general, on one hand, and for DMs specifically, on the other – through the design of a method that allows a contrastive analysis of DMs and (dis)fluencies across the *DisFrEn* and *Val.Es.Co. 2.0.* datasets. Annotations in these corpora differ in three main ways:

First, fluencemes are tagged in the *DisFrEn* corpus according to the protocol established in Crible et al. (2016) where fluencemes are annotated at word-level. In contrast, a discourse-level tagging approach is taken in the *Val.Es.Co. 2.0* corpus, in which a whole fluency structure named self-repair (see Schegloff et al. 1977 and Levelt 1983) and its constituent parts were annotated in a sample of conversations (see Pascual 2016), taking as a basis the Val.Es.Co. system of discourse units (Briz et al. 2002 and Briz and Grupo Val.Es.Co. 2014).

Second, for the functional annotation of DMs *DisFrEn* uses the interplay of four domains – ideational, rhetorical, sequential and interpersonal – and their associated functions (Crible under review), whereas in *Val.Es.Co. 2.0* the annotated discourse units of the Val.Es.Co. model – specifically the subact (Hidalgo and Padilla 2006) – provide cues to identify DMs, which are associated with three possible functions: text structuring, modality and interpersonality.

Third, syntactic and positional parameters of DMs are also annotated differently in both corpora. In the *DisFrEn* corpus, position is articulated in relation to three major units: the clause, the whole dependency structure and the turn. In *Val.Es.Co. 2.0*, the position of DMs is established hierarchically according to the interactional units of the discursive model: subact, act, intervention, turn, etc.

The results of this study are twofold. On the one hand, we provide an account of the methodological barriers that had to be overcome in order to carry out a corpus-based analysis. We offer a number of practical solutions to the problem of mapping diverse systems of annotation in two different corpora. For instance, we found that there is a correspondence between the four functional domains for DMs annotated in *DisFrEn* and the three functional types of subact of *Val.Es.Co. 2.0* (e.g. domains such as the interpersonal or the rhetorical one have a high correspondence with the interpersonal adjacent subact and with the modal adjacent subact, respectively). On the other hand, we identify patterns of uses of DMs within (dis)fluent structures in English, French and Spanish, highlighting the functional equivalences and non-equivalences among these three languages. The results of this study uncover recurring patterns of DM use that transcend crosslinguistic variation and that may therefore be regarded as potentially universal discourse 'constructions' (Fischer and Alm 2013). The study not only fills a gap in crosslinguistic fluency research, it seeks, moreover, to bring us closer to detecting universals of (dis)fluency.

Auer, P. (2009): On-line syntax: Thoughts on the temporality of spoken language. *Language Sciences*, 31, 1-13.

Briz, A. (1998): *El español coloquial en la conversación. Esbozo de pragmagramática*. Barcelona: Ariel.

Briz, A. and Grupo Val.Es.Co. (2002): Un sistema de unidades para el estudio del lenguaje coloquial. *Oralia*, 6, 7-61.

Briz, A. and Grupo Val.Es.Co. (2014): Las unidades del discurso oral. La propuesta

Val.Es.Co. de segmentación de la conversación (coloquial). *Estudios de Lingüística del Español*, 35, 1, 11-71 [online: http://infoling.org/elies/35/elies35.1-2.pdf].

Cabedo, A. y Pons, S. (eds.) (2013): *Corpus Val.Es.Co 2.0.* [online: http://www.valesco.es].

Crible, L. (under review). Modeling discourse markers and disfluency: Paradigmatic corpus annotation in *DisFrEn*. Submitted to *International Journal of Corpus Linguistics*.

Crible L., Dumont A., Grosman I., Notarrigo I. (2016). Annotation manual of fluency and disfluency markers in multilingual, multimodal, native and learner corpora. Version 2.0. Technical report. Université catholique de Louvain and Université de Namur.

Fischer, K. and Alm, M. (2013): A radical construction grammar perspective on the modal particle-discourse particle distinction. In L. Degand, B. Cornillie and P. Pietrandrea (eds), *Discourse Markers and Modal Particles: Categorization and Description*, Amsterdam: John Benjamins, 47-87.

Götz, S. (2013): *Fluency in Native and Nonnative English Speech*. Amsterdam: John Benjamins.

Hidalgo, A. and Padilla, X. (2006): Bases para el análisis de las unidades menores del discurso oral: los subactos. *Oralia*, 9, 109-143.

Levelt, W. (1893): Monitoring and self-repair in speech. *Cognition*, 14, 41-104.

Mann, W. C. and Thompson, S. A. (1988): Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8, 3, 243-281.

Pallaud, B., Rauzy, S., and Blâche, P. (2013): Auto-interruptions et disfluences en français parlé dans quatre corpus du CID. *TIPA Travaux Interdisciplinaires sur la Parole et le Langage*, 29.

Pascual, E. (2016): Annotating discourse units in spontaneous conversations: The challenge of self-repairs. Poster presented at the *Second Action Conference and MC Meeting: Cross-Linguistic Discourse Annotation: Towards Shared Resources, Tools and Methods* (TextLink: COST Action IS1312), Budapest (Karoli Gaspar University of the Reformed Church), 11–14 April, 2016.

Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., and Webber, B. (2008): The Penn Discourse TreeBank 2.0. In *Proceedings of the 6th Language Resources and Evaluation Conference (LREC'08),* 2961-2968.

Schegloff, E., Jefferson, G. and Sacks, H. (1977): The preference for self-correction in the organization of repair in conversation. *Language*, 53, 361-382.

Schiffrin, D. (1987): *Discourse Markers*. Cambridge: CUP.

Shriberg, E. (1994): *Preliminaries to a Theory of Speech Disfluencies* (Unpublished doctoral dissertation). University of California at Berkeley.

Disfluency markers as a potential tool for diagnosis, evaluation and monitoring in Primary Progressive Aphasia

PASTORIZA-DOMÍNGUEZ, P.[1]; DIÉGUEZ-VIDE, F.[1]; GÓMEZ-RUIZ, M.I.[2]; BELLO-LÓPEZ, J.[2] and Asunción ÁVILA-RIVERA[2]

Universitat de Barcelona[1], Neurology Unit, Hospital General de Hospitalet de Llobregat[2]

I. INTRODUCTION

PPA is a relatively recent clinical entity[1] whose description and diagnosis are still matter of debate in the realm of Neurology due to its pathological and clinical heterogeneity[2], its apparent overlap with other dementia types and its implications for cognitive and linguistic theory[3].From the clinical viewpoint, what was first considered a fluent vs. nonfluent dichotomy formed by PPA-S and PPA-G has led to a much more complex phenotypical picture after the current gold standard was established[4], with PPA-L being set apart from the nonfluent group and mapped into the PPA syndrome as a single entity. PPA-S displays anomia and occasional word comprehension problems. PPA-G exhibits effortful, telegraphic speech due to motor planning disruptions and/or frank agrammatism while PPA-L presents with a slow speech rate reflecting word-finding problems and phonemic paraphasias.

The definition of fluency has always been elusive in the Aphasiology field, as it refers to a multidimensional, composite measure derived from a number of features of speech production that may appear independently and are concurrently identified as the symptoms and the syndrome in cerebrovascular aphasia and PPA[4],[5]. What is more, in dementia fluency is subject to the progressive nature of the disease and is therefore dynamic from the longitudinal point of view[7]. Following Friederici's model[6],[7] these differential speech features are subserved by distinct language subsystems with different anatomic locations[5],[11]. Importantly, these subsystems are interconnected through large-scale extended multilayer neural networks[12], constrained by probabilistic factors that operate in parallel and which determine the final output after those are weighed out and integrated online according to task type, the nature and characteristics of the information decoded and encoded (e.g. word and structure frequency) and context[13],[14].

Disfluency typifying and assessment are key to our research, with the objective of disentangling articulatory and apraxic deviations from frank agrammatism in order to delineate and pinpoint the subdivision of the agrammatic/nonfluent variant into two etiological groups, as recent neuroimaging and biomedical findings seem to support[15]. This dichotomy would respond to damage in the motor and pre-motor areas in the case of nonfluent patients presenting with halting, articulatorily defective effortful speech, whereas it would be caused by deterioration in orbito-frontal areas involved in syntactic organization in the case of individuals displaying slow, hesitant telegraphic speech[16]. Disfluency evaluation is also central to the distinction of these subgroups from logopenic patients, a controversial group due to its neuropathological and clinical overlap with AD. These individuals' word-finding problems-allegedly caused by alterations in their phonological memory buffer[17]- could be mistaken for the anomic speech of Alzheimer patients[18]. In addition, the existence of a considerable number of unclassifiable cases[19]–[21] makes it necessary to attempt a description of relevant features for treatment and the potential reconsideration of the current gold standard[15],[22],[23].

II. METHODOLOGY

Our current study draws on transdisciplinary scientific and theoretical data in order to conduct a complete characterization of the linguistic and communicative profile of PPA, AD and MCI. This cohort choice responds to the necessity of describing the constellation of deficits that identify each group in order to find indexes of divergent (and/or convergent) performance for classification and early diagnosis, granted with the inclusion of probable pre-dementia stage MCI patients as well as age-matched healthy controls.

Samples of spontaneous speech in the form of narratives elicited from picture descriptions are analysed from a blind data-driven approach with Praat[24], creating a multi-level labelling and quantification system to account for performance at the phonemic/phonetic, lexico-semantic, morphologic, syntactic and pragmatic levels. Disfluency scrutiny is to the study, with a series of markers being subject to qualitative and quantitative analysis, namely repetitions, false starts, sequence repairs, abandoned structures and lengthening of vowel sounds. Their type, rate, location and distribution are analysed in order to attempt discrimination and group assignment, paying special attention to the position of disfluency markers within and out of phrase/sentence boundaries and to correlations between disfluency types and lexical features of affected and contiguous words such as grammatical category and frequency as well as phonetic/phonemic paraphasias and errors. The rate and distribution of filled and unfilled pauses in relation to such coordinates are also evaluated, all in order to disentangle disfluency caused by anomia due to semantic impairments (PPA-S), by lexical access and/or phonological planning issues (PPA-L) or by apraxic and/or motor problems and grammatical linearization disruptions (PPA-G).

III. CONCLUSION

The purpose of our study is to establish a sound description and framework for early detection and classification of PPA groups according to their linguistic and communicative performance, focusing on disfluency analysis in the first stages of the disease and, in the future, to sketch a longitudinal portrayal of the dissolution of the language system due to the neurodegenerative processes involved in PPA, acknowledging its progressive, changing nature. The methodology used will be explained in detail, outlining our preliminary results and interpreting their repercussions for current knowledge on PPA and related dementias.

ACKNOWLEDGMENT

This work is partially supported by La Caixa Foundation through a PhD grant awarded to P.P.D.

[1] M. M. Mesulam, *Slowly progressive aphasia without generalized dementia.*, vol. 11, no. 6. 1982, pp. 592–598.

[2] J. M. Harris and M. Jones, 'Pathology in Primary Progressive Aphasia Syndromes', *Curr. Neurol. Neurosci. Rep.*, vol. 14, no. 8, 2014.

[3] M. Harciarek and A. Kertesz, 'Primary progressive aphasias and their contribution to the contemporary knowledge about the brain-language relationship', *Neuropsychol. Rev.*, vol. 21, no. 3, pp. 271–287, 2011.

[4] M. L. Gorno-Tempini, a E. Hillis, S. Weintraub, a Kertesz, M. Mendez, S. F. Cappa, J. M. Ogar, J. D. Rohrer, S. Black, B. F. Boeve, F. Manes, N. F. Dronkers, R. Vandenberghe, K. Rascovsky, K. Patterson, B. L. Miller, D. S. Knopman, J. R. Hodges, M. M. Mesulam, and M. Grossman, 'Classification of primary progressive aphasia and its variants', *Neurology*, vol. 76, pp. 1–10, 2011.

[5] P. Feyereisen, A. Pillon, and M.-P. de Partz, ‘On the measures of fluency in the assessment of spontaneous speech production by aphasic subjects’, *Aphasiology*, vol. 5, no. 1, pp. 1–21, 1991.

[6] J. K. Gordon, ‘The fluency dimension in aphasia’, *Aphasiology*, vol. 12, no. 7, pp. 673– 688, 1998.

[7] M. Harciarek, E. J. Sitek, and A. Kertesz, ‘The patterns of progression in primary progressiva aphasia: Implications foor assessment and management’, *Aphasiolog*, vol. 28: 8-9, no. January 2015, pp. 964–980, 2014.

[8] A. D. Friederici, ‘The brain basis of language processing: From structure to function.’, *Physiol. Rev.*, vol. 91, no. 4, pp. 1357–92, 2011.

[9] A. D. Friederici, ‘The cortical language circuit: From auditory perception to sentence comprehension’, *Trends in Cognitive Sciences*, vol. 16, no. 5. pp. 262–268, 2012.

[10] S. M. Wilson, M. L. Henry, M. Besbris, J. M. Ogar, N. F. Dronkers, W. Jarrold, B. L. Miller, and M. L. Gorno-Tempini, ‘Connected speech production in three variants of primary progressive aphasia’, *Brain*, vol. 133, no. 7, pp. 2069–2088, 2010.

[11] C. E. Leyton and K. J. Ballard, ‘Primary progressive aphasia: conceptual evolution and challenges’, *Neurosci. Neuroeconomics*, vol. 5, pp. 9–18, 2016.

[12] J. L. Mcclelland, D. E. Rumelhart, and G. E. (eds. . Hinton, ‘The Appeal of Parallel Distributed Processing’, *Parallel Distrib. Process. Explor. Microstruct. Cogn.*, vol. I, pp. 3– 44, 1986.

[13] M. S. Seidenberg and M. C. MacDonald, ‘A Probabilistic Constraints Approach to Language Acquisition and Processing’, *Cogn. Sci.*, vol. 23, no. 4, pp. 569–588, 1999.

[14] N. Chater and C. D. Manning, ‘Probabilistic models of language processing and acquisition’, *Trends Cogn Sci*, vol. 10, no. 7, pp. 335–344, 2006.

[15] H. Botha, J. R. Duffy, J. L. Whitwell, E. A. Strand, M. M. Machulda, C. G. Schwarz, R. I. Reid, A. J. Spychalla, M. L. Senjem, D. T. Jones, V. Lowe, C. R. Jack, and K. A. Josephs, ‘Classification and clinicoradiologic features of primary progressive aphasia (PPA) and apraxia of speech’, *Cortex*, vol. 69, pp. 220–236, 2015.

[16] K. A. Josephs, J. R. Duffy, E. A. Strand, M. M. Machulda, M. L. Senjem, V. J. Lowe, C. R. Jack, and J. L. Whitwell, ‘Syndromes dominated by apraxia of speech show distinct characteristics from agrammatic PPA’, *Neurology*, vol. 81, no. 4, pp. 337–345, 2013.

[17] M. L. Gorno-Tempini, S. M. Brambati, V. Ginex, J. Ogar, N. F. Dronkers, A. Marcone, D. Perani, V. Garibotto, S. F. Cappa, and B. L. Miller, ‘The logopenic/phonological variant of primary progressive aphasia’, *Neurology*, vol. 71, no. 16, pp. 1227–1234, 2008.

[18] C. E. Leyton and J. R. Hodges, ‘Towards a clearer definition of logopenic progressive aphasia’, *Curr. Neurol. Neurosci. Rep.*, vol. 13, no. 11, 2013.

[19] S. A. Sajjadi, K. Patterson, and P. J. Nestor, ‘Logopenic, mixed, or Alzheimer-related aphasia?’, *Neurology*, vol. 82, no. 13, pp. 1127–1131, 2014.

[20] M.-M. Mesulam, C. Wieneke, C. Thompson, E. Rogalski, and S. Weintraub, ‘Quantitative classification of primary progressive aphasia at early and mild impairment stages’, *Brain*, vol. 135, no. Pt 5, pp. 1537–53, 2012.

[21] M. R. Wicklund, J. R. Duffy, E. a. Strand, M. M. Machulda, J. L. Whitwell, and K. a. Josephs, ‘Quantitative application of the primary progressive aphasia consensus criteria’, *Neurology*, vol. 82, no. 13, pp. 1119–1126, 2014.

[22] M. M. Mesulam and S. Weintraub, ‘Is it time to revisit the classification guidelines for primary progressive aphasia?’, *Neurology*, vol. 82, no. 13, pp. 1108–1109, 2014.

[23] C. E. Leyton and J. R. Hodges, 'Differential diagnosis of primary progressiva aphasia variants using the international criteria', *Aphasiology*, vol. 28: 8-9, no. October 2013, pp. 909–921, 2014.

[24] P. Boersma and D. Weenink, 'Praat: doing phonetics by computer (Version 4.5.)[Computer program]', *Retrieved from http//www.praat.org/*, vol. 5, no. 9/10, pp. 341–345, 2007.

Finnish learners' speech fluency in L2 English: Exploring fluency in interaction and L1 speaking style

Pauliina PELTONEN

University of Turku

L2 speech fluency has traditionally been approached from a psycholinguistic perspective, being viewed as a result of smoothly running, underlying cognitive speech production processes (e.g. Lennon 2000: 26). Empirical approaches to fluency have applied this view by focusing on the temporal aspects of speech (i.e. the amount of speech and the extent of pausing) that reflect processing efficiency. The psycholinguistic approach is also evident in the types of speech samples that are examined: they are predominantly monologues. Studies following this tradition have provided important information about the most reliable quantitative fluency measures in monologue conditions. Speech rate, articulation rate, mean length of run and pause location have been found to differentiate between fluent and disfluent speakers and learners from native speakers (e.g. Kormos & Dénes 2004, Hilton 2014, Kahng 2014). However, the results regarding possibly multifunctional features, such as filled pauses (FPs) and repetitions, are less clear and require further study.

The lack of between-groups differences and high within-group variations demonstrated in earlier studies (e.g. Kahng 2014, Peltonen & Lintunen 2016) point to individual differences in the use of FPs and repetitions. Despite often being referred to as "disfluency markers", FPs and repetitions can even be considered fluency-enhancing, since they reduce time spent in silence and can be regarded as solutions to the processing time pressure associated with speech production (cf. Götz' 2013 view of disfluency markers as "speech management strategies"). Therefore, approaching FPs and repetitions as problemsolving mechanisms rather than as features contributing to disfluency is essential for advancing our knowledge of how learners maintain fluency with limited skills in an L2.

In the oral presentation, I will discuss the main results of two interrelated substudies that were conducted as a part of a larger project exploring Finnish learners' speech fluency in L2 English from different perspectives. The project addressed two main aspects that have been identified as essential for further study (e.g. Wright & Tavakoli 2016), but have so far not been extensively studied: L2 fluency in a dialogue setting, including examination of interactional fluency (Study 1) and the effect of individual L1 speaking style on L2 speech fluency (Study 2).

In addition to examining fluency in the relatively unexplored dialogic mode (but see the pioneering study by Riggenbach 1991; more recently Witton-Davies 2014, Tavakoli 2016), Study 1 focused on examining problem-solving mechanisms along with established temporal

fluency measures. In the study, two main types of problem-solving mechanisms (referred to together as fluency resources) were regarded as potentially contributing to fluency: *stalling mechanisms* (e.g. FPs and repetitions) that help in coping with processing time pressure and *communication strategies* (CSs, see e.g. Dörnyei & Scott 1997) that aid in overcoming mainly lexis-related problems (see Dörnyei & Kormos 1998). Even though references to strategic competence and L2 speech fluency have been made in theory (e.g. Segalowitz 2010), the link between the two has previously not been studied empirically (but see Dörnyei 1995 on the effects of teaching stalling mechanisms and CSs on speech rate).

In Study 2, the focus was on examining the effect of individual L1 speaking style on L2 speech. Few studies so far have included both L1 and L2 samples from the same subjects (but see Raupach 1980; more recently De Jong et al. 2015), though there is some evidence to suggest that at least silent pause duration is more indicative of L1 speaking style than L2 proficiency (De Jong et al. 2015). Also the above-mentioned individual differences associated with the use of stalling mechanisms may be related to speaking style.

The participants in both studies were 42 Finnish learners of English at two school levels (16 ninth graders, 15-year-olds, and 26 second year upper secondary school students, 17–18-year-olds). The data for Study 1 consist of problem-solving tasks that the subjects conducted in pairs. Their task was to rank 16 items (pictures) in the order of importance for survival on a desert island. The data for Study 2 include comparable picture description monologues in Finnish (L1) and English (L2) from the same subjects.

In both studies, subjects from different school levels were compared for their temporal fluency (speech rate, articulation rate, mean length of run as well as the frequency, length and location of silent pauses) and stalling mechanisms (filled pauses, drawls, fillers and repetitions) to find out how learners of English from two school levels differed in their fluency. Study 1 also included dialogue-specific measures (the number and duration of turn pauses as well as other-repetitions and collaborative completions) and communication strategies. Both studies employed a mixed methods approach to fluency analysis, complementing the quantitative results (groupwise comparisons of the frequencies of fluencyrelated features) with a qualitative analysis. Following earlier mixed methods studies (e.g. Towell, Hawkins & Bazergui 1996; Brand & Götz 2011; Hilton 2014), the qualitative component focused on examining the fluency-related features for other aspects besides their frequency, especially the functions of stalling mechanisms in context to account for their multifunctionality, and relating the findings to individual learners' fluency profiles.

Preliminary results for the studies suggest that in the problem-solving task, upper secondary school students used more fillers and repetitions than ninth graders to facilitate fluency. In the L2 English monologues, in addition to fillers and repetitions, also drawls differentiated the groups in the quantitative analysis. For temporal measures, the differences were clear in the dialogue: the upper secondary school subjects outperformed the ninth grade group on 11 of the 13 temporal measures. In the L2 English monologues, the differences were minor, since differences only on the mean length of run were statistically significant. The qualitative analysis of Study 1 indicated that combinations of different types of CSs, as well as combinations of stalling mechanisms and CSs, were common. Furthermore, stalling mechanisms helped in keeping the flow of talk going during difficulties and therefore compensated for fluency losses caused by the higher processing load associated with communication strategies.

Brand Christiane & Götz Sandra 2011. Fluency versus accuracy in advanced spoken learner language: A multi-method approach. *International Journal of Corpus Linguistics* 16: 255–275.

De Jong Nivja H., Groenhout Rachel, Schoonen Rob & Hulstijn Jan J. 2015. Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics* 36: 223–243.

Dörnyei Zoltán 1995. On the teachability of communication strategies. *TESOL Quarterly* 29: 55–85.

Dörnyei Zoltán & Kormos Judit 1998. Problem-solving mechanisms in L2 communication: A psycholinguistic perspective. *Studies in Second Language Acquisition* 20: 349–385.

Dörnyei Zoltán & Scott Mary Lee 1997. Communication strategies in a second language: Definitions and taxonomies. *Language Learning* 47: 173–210.

Götz Sandra 2013. *Fluency in Native and Nonnative English Speech*. Amsterdam & Philadelphia: John Benjamins.

Hilton Heather 2014. Oral fluency and spoken proficiency: Considerations for research and testing. In Pascale Leclercq, Amanda Edmonds & Heather Hilton (eds). *Measuring L2 Proficiency: Perspectives from SLA*, 27–53. Bristol, UK: Multilingual Matters.

Kahng Jimin 2014. Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning* 64: 809–854.

Kormos Judit & Dénes Mariann 2004. Exploring measures and perceptions of fluency in the speech of second language learners. *System* 32: 145–164.

Lennon Paul 2000. The lexical element in spoken second language fluency. In Heidi Riggenbach (ed.). *Perspectives on Fluency*, 25–42. Ann Arbor, MI: University of Michigan.

Peltonen Pauliina & Lintunen Pekka 2016. Integrating quantitative and qualitative approaches in L2 fluency analysis: A study of Finnish-speaking and Swedish-speaking learners of English at two school levels. *European Journal of Applied Linguistics* 4: 209–238.

Raupach Manfred 1980. Temporal variables in first and second language speech production. In Hans W. Dechert & Manfred Raupach 1980 (eds). *Temporal Variables in Speech*, 263–270 . The Hague: Mouton.

Riggenbach Heidi 1991. Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes* 14 (4): 423–441.

Segalowitz Norman 2010. *The Cognitive Bases of Second Language Fluency*. New York: Routledge.

Tavakoli Parvaneh 2016. Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching* 54: 133–150.

Towell R., Hawkins R. & Bazergui N. 1996. The development of fluency in advanced learners of French. *Applied Linguistics* 17: 84–119.

Witton-Davies Giles 2014. The Study of Fluency and its Development in Monologue and Dialogue. Unpublished PhD dissertation. Lancaster University, UK.

Wright Clare & Tavakoli Parvaneh 2016. New directions and developments in defining, analyzing and measuring L2 speech fluency. *International Review of Applied Linguistics in Language Teaching* 54: 73–77.

Morphosyntactic regularities in retracing phenomena in the speech of second language learners of French

Hugues PETERS

UNSW, Australia

Corpus studies have shown that native speakers' spontaneous speech contains disfluencies, such as pauses, filled pauses, and various forms of retracing (Candéa 2000, Blanche-Benveniste 2010). Such behaviour is accounted for by the cognitive load required in working memory for the activation of words and the processing of sentences in real time (Levelt 1989). Unsurprisingly, similar phenomena characterize the spontaneous speech of L2 learners, who, depending on their level of proficiency, produce them at significantly higher rates than native speakers (Towell *et al*. 1996, Temple 2000, Hilton 2014).

The following utterance provides an example of disfluent French L2 speech:

(1) File: UWI Corpus: 0831_Ex.cha (l.177-183):

*STU: et <il y a un> [/-] &hum le gagnant &hum (.) yeah@s &euh peut [//] va &euh (.) faire un &hum (.) cd avec leurs chansons, quelque chose comme ça, &*INV:oui, &euh oui .

%mor: conj|et det|le&MASC&SING n|gagnant L2|yeah v:mdl|aller&PRES&3SV v:mdllex|faire-INF det|un&MASC&SING n|cd&MASC prep|avec det:poss|leurs&PL n|chanson-PL&FEM det:gen|quelque n|chose prep|comme pro:dem|ça adv:yn|oui .

"and there is a (filler) the winner (filler) yeah (filler) can is-going-to (filler) make a (filler) cd with their songs, or something like that, (listener backchannel), (filler) yes."

Research in this domain has usually been linked to the establishment of quantifiable, objective, and reliable temporal measures correlated with the perception of fluency (Lennon 1990, Freed 2000). Tavakoli & Skehan (2005) distinguish three types of fluency: speed fluency, concerned with the rate of speech, breakdown fluency, concerned with silent and filled pauses, and repair fluency, concerned with repairs. Among the latter, Olynyk *et al*. (1990) establish a distinction between progressive repairs (repeats, but also fillers) that are not necessarily detrimental to fluency as they allow native speakers and highly proficient L2 learners to avoid silent pauses within utterances, and regressive repairs (self-repairs and false starts) that affect fluency.

Per current models of L2 speech production (Kormos 2006), the abundance of disfluencies in non-native speech is accounted for by a lack of automatization of the processes of lexical retrieval and syntactic production. It therefore involves the conscious arrangement of individual lexemes into syntactic units, a laborious process that rapidly overloads the working memory (Hilton 2014). Additionally, as suggested by Temple (2000), self-repairs of L2 learners differ not only quantitatively, but also qualitatively from those of native speakers: the former, again with variations depending on their proficiency, make many corrections of syntax and morphology, while the latter focus on lexis or register.

In the spirit of the grammatical study of self-repairs by Fornel & Marandin (1996), the present study investigates the morphosyntactic properties of retracing phenomena (repeats, self-repairs and false starts) produced by L2 learners of French. It is based on the careful quantitative

analysis of data from the UWI learner corpus (Péters 2014), and offers a qualitative analysis based on current linguistic theories, such as Distributed Morphology (Halle & Marantz 1993). Our hypothesis is that these phenomena, usually considered prime examples of performance hiccups, manifest in fact syntactic regularities, which can reveal aspects of the underlying grammatical competence of learners.

The UWI learner corpus is a longitudinal corpus of oral productions by 10 Jamaican learners of French, engaging in 67 one-on-one conversations with the investigator over a nineteen-month period. This corpus of 54,824 words has been transcribed, annotated, and encoded using the CHILDES conventions (MacWhinney 2000). Each orthographic line is coupled with a morphological decomposition line, pruned of fillers and retracing, as shown in (1) above, and has been linked with the corresponding audio-file in electronic format. The CLAN system provides inbuilt tools for the automatic analysis of the corpus.

The participants, all diglossic natives of Jamaican Creole and Jamaican English, were learning French at the University of the West Indies, and had never been immersed in a French speaking country at the time of the interviews. The corpus was primarily compiled to assess the grammatical characteristics of their Interlanguage, and especially the influence of their L1s on the acquisition of an L2. However, issues around the transcription of disfluencies immediately came to the fore as the transcription conventions in CHILDES provide for researchers to encode retracing in such a way that the original segment is not parsed on the morphological tier.

Our exclusive concern is on the analysis of unexpected asymmetries characterising retracing of morphosyntactic features by learners at different levels of proficiency: more precisely, articles and verbal clitics, two important topics of SLA research on the acquisition of French (Prévost & Paradis 2004).

In confirmation of previous research showing the prominence of the masculine gender in L2 interlanguage grammar of definite articles (Dewaele & Véronique 2001, Granfeldt 2004, Péters & Stewart 2009), the data shows that, in cases of self-repair with a change of gender of a definite determiner, most occurrences go from the masculine *le* to the feminine *la*. So, self-repairs of the type: *le [//] la cité* (the city) are observed, but very rarely of the type: *la [//] le cours* (the course). In this case, if lexical activation proves difficult, repeats: *le [/] le mot* (the word) or *la [/] la présence* (the presence), or repairs involving the whole noun phrase: *<la droit> [//] le droit* (law studies) are observed. Another feature, more characteristic of lower proficiency students, that clearly distinguishes them from native speakers, is the tendency to omit the third person subject pronoun when repeating lexical verbs (*ils enseignent [/] enseignent*). Interestingly, they never omit that pronoun with auxiliary verbs, except in cases of negative or interrogative structures, and never omit the first-person pronoun. These facts inform us on the featural composition given to clitic pronouns and on the default value of masculine definite determiners in the Interlanguage competence of these speakers.

Our goal will be to propose a syntactic representation of these retracing strategies. Such quantitative and qualitative research can have benefits not only for a better understanding of the conscious (Monitor) and unconscious grammatical processes, but also to inform the pedagogy of French as a second or foreign language. Further research will compare these results with other L2 learners of French from other linguistic background and with native speakers.

Blanche-Benveniste C. (2010). *Approches de la langue parlée en français*. Paris: Ophrys.

Candéa M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané*. Thèse de doctorat nouveau régime, Université de Paris III.

Dewaele, J.-M. & Véronique, D. (2001). Gender assignment and gender agreement in advanced French Interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition* 4, 275–297.

Fornel M. de & Marandin J.-M. (1996). L'analyse grammaticale des auto-réparations. *Le gré des langues*, 10, 8-68.

Freed B. (2000). Is fluency, like beauty, in the eyes and ears of the beholder? In H. Riggenbach (ed.), *Perspectives on fluency* (pp.243–265). Ann Arbor, MI: University of Michigan Press.

Granfeldt J. (2004). Domaines syntaxiques et acquisition du français langue étrangère. L'exemple du DP. *Acquisition et Interaction en Langue Etrangère*, 20, 47-84.

Halle M. & Marantz A. (1993). Distributed Morphology and the Pieces of Inflection. In: K. Hale & S.J. Keyser (eds.), *The View from Building 20* (pp.111–176). Cambridge, MA: MIT Press.

Hilton H. (2014). Oral fluency and spoken proficiency: considerations for research and testing. In: P.

Leclerq, A. Edmands & H. Hilton (eds.), *Measuring L2 proficiency. Perspectives from SLA* (pp.2751). Bristol: Multilingual Matters.

Kormos J. (2006). Speech production and second language acquisition. Mahwah, NJ: Lawrence Erlbaum.

Lennon P. (1990). Investigating fluency in EFL. *Language Learning*, 40 (3), 387-417.

Levelt W. (1989). *Speaking: from intention to articulation.* Cambridge, MA: MIT Press.

MacWhinney B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Mahwah, NJ: Lawrence Erlbaum Associates.

Olynyk M., D'Anglejan A. & Sankoff D. (1990). A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. In: R. Scarcella, R. Anderson & S. Krashen (eds.), *Developing communicative competence in a second language* (pp.139-155). Rowley, MA: Newbury House.

Péters H. (2014). Étude sur la complémentation infinitive des verbes de déplacement sur la base d'un corpus de productions orales d'apprenants jamaïcains de FLE et suggestions pédagogiques. In: L. de Serres, F. Ghillebaert, P-A. Mather & A. Bosch (eds.), *Aspects culturels, linguistiques et didactiques de l'enseignement-apprentissage du français à un public non-francophone* (pp.62– 84). Québec: Association Internationale des Études Québécoises.

Péters H. & Stewart M. (2009). The syntax of determination in a corpus of oral productions of Jamaican learners of French: word order and agreement patterns in nominal phrases. *The International Journal of the Humanities*, 7, 51-62

Prévost P. & Paradis J. (eds.). (2004). *The acquisition of French in different contexts. Focus on functional categories.* John Benjamins.

Tavakoli P. & Skehan P. (2005). Strategic planning, task structure and performance testing. In: R. Ellis (ed.), *Planning and task performance in a second language* (pp.239–277). Amsterdam: John Benjamins.

Temple L. (2000). Second language learner speech production. *Studia Linguistica*, 54 (2), 288-297. Towell R., Hawkins R. & Bazergui N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, 17(1), 84–119.

Pauses in connected speech of early Alzheimer patients: comparison of a picture-based and a memory-based discourse

Aurélie PISTONO[1,2], Jérémie PARIENTE[1,3], Catherine BÉZY[3], Béatrice LEMESLE[3], Bérengère PAGES[3], Johanne LE MEN[3] & Mélanie JUCLA[2]

Toulouse University[1], University of Toulouse II-Jean Jaurès[2], Toulouse University Hospital[3]

Background

Although episodic memory impairment is often observed in Alzheimer's disease (AD), language difficulty is one of the deficits that is frequently reported in patients. Research has shown that lexicalsemantic deficits can occur in the very earliest stage of AD (Amieva et al., 2008). These deficits have been observed in fluency (Clark et al., 2009), naming (Domoto-Reilly et al., 2012), and semantic tasks (Joubert et al., 2010). More and more studies have been focusing on discourse production in AD, with a view to assessing the functional use of language and cognition. Some have focused on pause production, revealing that patients make extensive use of pauses during speech (Singh et al., 2001; Gayraud et al., 2011). This has been attributed to lexical retrieval difficulties, but pausing may also reflect other forms of cognitive processes. In a previous work, authors showed that pausing behaviour was positively correlated with anterograde episodic memory, when patients produced an autobiographical discourse. This result implies that pausing may also constitute a compensatory mechanism used to improve mental time travel and autobiographical discourses' completeness. In the current work, pausing within two types of narratives was investigated: within a personal memory-based narrative and a fictive picture-based narrative. We aimed to address the following questions: do patients with AD produce longer and/or more frequent pauses in both types of narratives? Are there different "pausing" pattern depending the type of narrative (memory-based vs. picture-based)? Does "pausing" reflect different cognitive processing in both groups of participants?

Methods

17 patients suffering from prodromal AD (Dubois et al., 2014) and 17 matched healthy controls were recruited. Each participant underwent a neuropsychological assessment and a language assessment. It included two narrative discourses: a picture-based narrative, made up of 5 sequential pictures and a memory-based one: recalling a specific event that occurred with the experimenter during the assessment.

The memory-based narrative was created specifically for the study and followed the same structure as the picture-based narrative: initial state; perturbation; event itself and final state. An episodic memory scoring was also attributed to the recall according to a specific grid. It allows an ecological episodic memory assessment.

Both narratives were recorded and orthographically transcribed with the CLAN program, using CHAT conventions (MacWhinney, 2011). Disfluencies were annotated, among which pauses. A pause was defined as a silence or hesitation starting from 200 ms (Gayraud et al., 2011) and marked using PRAAT.

Four measures were used:

- Number of words, including well-formed words as repeated words and rephrasing but excluding false starts (Marini et al., 2011, 2012);

- Speech rate, measured using the number of words/total speech duration in seconds (Marini et al., 2011, 2012);
- Pause frequency, normalized per 100 words;
- Pause length, measured using the median of each participant.

Statistical analyses: As regards the memory-based narrative, a Cohen's Kappa was calculated on memory scoring between an experimenter and a speech therapist, for all narratives.

As for narratives assessment, controls' and patients' narratives were first compared using the nonparametric Mann Whitney test. Then, both narratives were compared within each group using the non-parametric Wilcoxon test. Finally, pause frequency and pause length within each group and each discourses were correlated with cognitive performance with non-parametric Spearman correlations: global cognition (Mini Mental State Examination, MMSE), anterograde memory (Doors and people test; Free and cued selective reminding test), processing speed (Trail Making Test, TMT A), mental flexibility (TMT B-A), verbal fluency (fruits, verbs, letter V), confrontation naming (objects, action, celebrities), syntactic comprehension, semantic verification.

Results

Cohen's Kappa for memory scoring was 0.86, which reflects a high inter-judge reliability.

Intergroup comparisons: Regarding the pictured-based narrative, patients did not produce fewer words than healthy controls but had lower speech rates ($p<0.01$). Indeed, they produced longer ($p=0.01$) and more frequent pauses ($p<0.01$). Concerning the memory-based narrative, patients did not produce fewer words either. However, they got lower speech rates ($p<0.05$) with longer pauses ($p<0.01$) but not more frequent ones.

Correlations: In the healthy group, there was no correlation between pause length and frequency with cognitive performance. In the patients' group, pausing was correlated with various cognitive processes depending on the narrative task. Within the pictured-based narrative, pause length was negatively correlated with verbal fluency (fruits, verbs, letter V) and positively correlated with longer processing speed (TMT A). There was no significant correlation with pause frequency. Within the memory-based narrative, pauses frequency was positively correlated with the recall itself, but also with verbal (Free and Cued Selective Reminding Test) and visual (Doors and people Test) anterograde memory, while no correlation with pause length was observed.

Discussion

The use of two types of narratives evidenced two major results. First, a memory-based narrative involves less pausing amongst healthy people, probably because it reflects real life discourses. On the contrary, a picture-based sequential narrative is unusual and more constraint, which leads to an increase of the number of pauses compared to a memory-based narrative. As the ability for recalling a personal event is impaired at the earliest stage of AD, the memory-based narrative is also disfluent in this group. Moreover, using two types of narratives, we put forward the influence of various cognitive loads during patients' discourse processing. Indeed, the pictured-based narrative probably placed cognitive resources upon discourse planning and executive functioning, which lead to the need for longer pauses. Patients who are less efficient in processing speed and lexical access make the longer pauses in this type of discourse. On the

contrary within a memory-based narrative, patients did not produce more pauses than healthy participants but used them differently. Pauses are more used in patients with the best mnemonic skills. We assume the pauses in this type of discourse may play as a compensatory mechanism used to improve recall.

Amieva, H., Le Goff, M., Millet, X., Orgogozo, J.M., Peres, K., Barberger-Gateau, P., et al. (2008). Prodromal Alzheimer's disease: successive emergence of the clinical symptoms. *Ann Neurol*; 64, 492–498.

Clark, L.J., Gatz, M., Zheng, L., Chen, Y.L., McCleary, C. & Mack, W. J. (2009). Longitudinal verbal fluency in normal aging, preclinical and prevalent Alzheimer disease. *American journal of Alzheimer's disease and others dementias*. 24 (6), 461-468.

Domoto-Reilly, K. (2012). Naming impairment in Alzheimer's disease is associated with left anterior temporal lobe atrophy. *Neuroimage*. 63 (348).

Dubois B, Feldman HH, Jacova C, … Cummings JL. (2014). Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 crite-ria. *Lancet Neurol*. 13, 614-629.

Gayraud, F., Lee, H-R., Barkat-Defradas, M. (2011). Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clinical Linguistics & Phonetics*. 25 (3), 198–209.

Joubert S., Brambati S.M., Ansado J., Barbeau E.J., Felician O., Didic M., Lacombe J., Goldstein R., Chayer C. & Kergoat M.J. (2010). The cognitive and neural expression of semantic memory impairment in mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia*. 48, 978-988.

MacWhinney, B. (2011). The Child Language Data Exchange System project : tools for analyzing talk. Electronic Edition. http://childes.psy.cmu.edu

Marini, A. et al., 2011. Narrative language in traumatic brain injury. Neuropsychologia, 49(10), pp.2904–10.

Marini, A., & Urgesi, C. (2012). Please get to the point! A cortical correlate of linguistic informativeness. Journal of Cognitive Neuroscience, 24(11), 2211–22. http://doi.org/10.1162/jocn_a_00283

Singh S, Bucks RS, Cuerden JM. (2001). Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech. *Aphasiology.* 15, 571-583.

Towards a comprehensive notion of fluency in simultaneous interpreting: An interdisciplinary approach

E. Macarena PRADAS MACÍAS

University of Granada

Interpreting fluency has been investigated from a number of perspectives: the effect of source-speech pauses on the interpreter (e.g. Barik 1973; Tissi 2000); the interpreter's pauses (e.g. Alexieva 1988; Shlesinger 1994; Mead 2000); the impact of silent pauses on the users' assessment of interpreting quality (Ahrens 1998; Lenglet 2014), and the users' perception of the fluency of the interpreter's delivery (Christodoulides & Lenglet 2014; Pradas Macías 2003b). Yet, it is still unclear which characteristics of pauses and other measurable features of speech make recipients perceive it as fluent (e.g. Rennert 2010).

The research presented here was carried out at the University of Granada in a number of research projects, the latest one being *Quality in Interpreting and Nonverbal Aspects* – QINV

for short.[6] QINV addresses both the objective features of discourse and the recipients' subjective impressions.

Within the "objective" line of our earlier research, we found similar silent-pause patterns in the production of different interpreters (Pradas Macías 2003a, 2009). They can be both a symptom of the interpreter's encoding difficulties or a sign that he or she is trying to help the listener to decode the message by following certain prosodic patterns, which may be specific to individual discourse types. The analysis of silent pause patterns involves a variety of research questions, for instance regarding the relation between lexical phenomena and the occurrence of silent pauses. Indeed, further research suggested that certain lexical categories in the source speech tend to trigger pauses in the target speech (Pradas Macías 2015).

In our "subjective" line of research, we learned that low fluency impinges on the overall perception of interpreting quality. In one study (Pradas Macías 2003b), an increased presence of silent pauses in interpreted speech led both interpretation users and interpreters to rate the interpretation more negatively, both in terms of fluency and of overall quality. In another study (Pradas Macías 2007), ratings were also impinged by disfluency symptoms added to an interpretation, such as filled pauses, repetitions and self-repairs. Moreover, we found some evidence of interpreters' using a *sui generis* type of silent-pause pattern that might give the listener a cue to distinguish interpreted from non-interpreted speech (Pradas Macías 2009). Furthermore, even when *sui generis* pauses are compressed to a residual duration, subjects may still be able to detect them (Pradas Macías 2015). This confirms earlier findings that perceived pauses do not necessarily match real pauses, putting in question the use of standard thresholds in the definition and the experimental treatment of pauses.

In QINV, our ongoing research project, we address both linguistic and extralinguistic aspects related to fluency. Its principal aim is to study the effect of acceptability and social perception on quality evaluation (García Becerra 2012). Its starting point is that acceptability may depend on users' expectations regarding the distinctive features of this type of speech and on their beliefs about the interpreter's role and visibility. The project puts the following hypothesis linked to fluency: 1) Interpreting-quality criteria such as fluency are ill defined and lead to a concept variability that can't lead to perception and assessment guidelines; 2) It is possible to outline some perception patterns, e.g., the relation between subjects' perception of fluency and the impression formed about the interpreter; 3) Fluency manipulation in the source speech has different impact on the interpreter's delivery; 4) Fluency manipulation in the interpreting product affects user's perception of quality; 5) Some prosody features interfere in the way the manipulations (may) affect more than one quality criterion.

The QINV project includes both observational and experimental studies. Raw and manipulated speeches and interpretations serve as material and the subjects ares experts and potential users. It is being carried out by a multidisciplinary team carrying out three strands of research. The first one follows in the steps of "vertical" studies in interpreting quality, which go beyond the "horizontal" distinction between quality criteria (such as fidelity, cohesion, intonation, and fluency) by exploring the interaction between specific aspects of those criteria. In particular, this strand examines the impact of specific prosodic aspects of fluency on users' impressions of overall quality and of other quality criteria. The second strand of research

6 Funded by the Spanish Ministry of Economy and Competitiveness, ref. FFI2014-56617-P. See http://qinv.ugr.es

contributes to the methodology for survey-based quality studies, aiming to advance questionnaire design and administration methods. The third strand is devoted to the preparation of experimental material. It explores options of experimental treatment for features of spoken discourse such as pauses which may have an impact on the perception of individual quality criteria.

Ahrens, Barbara. 1998. Nonverbale Phänomene und Belastung beim Konsekutivdolmetschen. In: TEXTconTEXT 1 (3-4), 213-234.

Alexieva, Bistra. 1988. Analysis of the simultaneous interpreter's output. In Paul Nekeman (ed.), Translation, Our Future. XI FIT World Congress Proceedings. Maastricht: Euroterm, 484488.

Barik, Henri. 1973. Simultaneous interpretation: Temporal and quantitative data. Language and Speech 16, 237-270.

Christodoulides, George & Lenglet, Cédric. 2014. Prosodic correlates of perceived quality and fluency in simultaneous interpreting. Speech Prosody 7 (Dublin, Ireland, from 20/05/2014 to 23/05/2014). In Nick Campbell, Dafydd Gibbon & Daniel Hirst (eds.); Social and Linguistic Speech Prosody. Proceedings of the 7th International Conference on Speech Prosody, 10021006. http://hdl.handle.net/2078.1/145308.

García Becerra, Olalla. 2012. La incidencia de las primeras impresiones en la evaluación de la calidad de la interpretación simultánea: un estudio empírico. Granada, Spain: Universidad de Granada PhD thesis. http://hera.ugr.es/tesisugr/21620234.pdf (1 October, 2013).

Gile, Daniel. 2000. The history of research into conference interpreting: A scientometric approach. Target 12 (2), 297-321.

Lenglet, Çedric. 2013. The Impact of Simultaneous Interpreting Prosody on Speech Perception: Proposal for an Experiment. In Gabriel González Núñez, Yasmine Khaled & Tanya Voinova (eds.), 2013. Emerging Research in Translation Studies: Selected Papers of the CETRA Research Summer School 2012.

Mead, Peter. 2000. Control of pauses by trainee interpreters in their A and B languages, The Interpreter's Newsletter 10. 89-102.

Pazos-Bretaña, José-Manuel, Olalla García Becerra & Rafael Barranco-Droege. 2010. Diseño y técnicas de explotación de un corpus oral para el análisis de parámetros de calidad en interpretación. In Isabel Moskowich-Spiegel Fandiño, Begoña Crespo García, Inés Lareo Martín & Paula Lojo Sandino (eds.), Language windowing through corpora = Visualización del lenguaje a través de corpus, 627–640. A Coruña, Spain: Universidade da Coruña. http://hdl.handle.net/10481/41022.

Pradas Macías, E. Macarena. 2015. User awareness of sui generis pause patterns in simultaneous interpreting. In Cornelia Zwischenberger & Martina Behr (eds.), Interpreting quality: A look around and ahead, 67–87. Berlin, Germany: Frank & Timme.

Pradas Macías, E. Macarena. 2009. Identificación del patrón pausístico para la medición de la calidad en interpretación simultánea. In Gerd Wotjak, Vessela Ivanova & Encarnación Tabares Plasencia (eds.), Translatione via facienda: Festschrift für Christiane Nord zum 65. Geburtstag, 235–252. (Studien Zur Romanischen Sprachwissenschaft Und Interkulturellen Kommunikation). Frankfurt am Main, Germany: Peter Lang. (5 August, 2013).

Pradas Macías, E. Macarena. 2007. La incidencia del parámetro fluidez. In Ángela Collados Aís, E. Macarena Pradas Macías, Elisabeth Stévaux & Olalla García Becerra (eds.), La evaluación de la calidad en interpretación simultánea: parámetros de incidencia, 53–70. Granada, Spain: Comares.

Pradas Macías, E. Macarena. 2006. Probing quality criteria in simultaneous interpreting: The role of silent pauses in fluency. Interpreting 8 (1). 25–43.

Pradas Macías, E. Macarena. 2003a. Análisis de la fluidez como parámetro de calidad en interpretación. In Ángela Collados Aís, María Manuela Fernández Sánchez & Daniel Gile (eds.), La evaluación de la calidad en interpretación: investigación, 97–106. Granada, Spain: Comares.

Pradas Macías, E. Macarena. 2003b. Repercusión del parámetro pausas silenciosas en la fluidez. Granada, Spain: Universidad de Granada.

Rennert, Sylvi. 2010. The impact of fluency on the subjective assessment of interpreting quality. The Interpreters' Newsletter 15, 101-115.

Shlesinger, Miriam. 1994. Intonation in the production and perception of simultaneous interpretation. In Sylvie Lambert & Barbara Moser-Mercer (eds.), Bridging the gap. Empirical research in simultaneous interpretation. Amsterdam/Philadelphia: John Benjamins, 225-236.

Tissi, Benedetta. 2000. Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis. The Interpreters' Newsletter 10, 103-127.

Fluency in Aviation English – the path to its description

Malila CARVALHO DE ALMEIDA PRADO

University of São Paulo

After a series of events which occurred due to lack of proficiency, the International Civil Aviation Organization (ICAO) implemented a new regulation to member states worldwide: in order to be accredited to operate internationally, pilots and air traffic controllers are required to undergo an English proficiency assessment. Governments must now manage a test (either produced "in house" or hired by institutions) which aims to assess the candidate's proficiency according to six different linguistic criteria: pronunciation, structure, vocabulary, fluency, comprehension and interaction, in six different levels. Although the requirements are described in an official document, DOC 9835 (ICAO 2010), which sheds light on practices for training and testing, there have been inconsistencies among exams all over the world (Alderson 2008; Emery 2014), indicating a need for more empirical research especially in the definition of the language to be evaluated (Bachman & Palmer 2010).

Throughout DOC 9835, ICAO emphasizes that the focus of the proficiency requirements must be on the plain language, that is, the language used in abnormal situations comprising elements other than Aeronautical Phraseology – a sublanguage highly conventionalized and documented. However, it is unarguably difficult to establish a precise definition to this scope of language (Moder & Halleck 2009) and, even though there has been growing research on the topic (Lopez et al. 2013), this problem still needs catering for.

The overall purpose of the present research is to fulfill an existing gap between real life communications and the constructs presented in the abovementioned document. A corpus of nearly 100 exchanges between pilots and ATCs in abnormal situations (conditions in which they produce language other than Aeronautical Phraseology) was compiled to this end. By means of corpus linguistics concepts and methodology, this paper outlines a segment of the research by demonstrating some of the investigation related to two of the linguistic criteria to be evaluated - fluency and interaction - in pursuance of a better understanding of what in fact occurs in radiotelephony communications and what makes the speakers fluent or not. The model of transcription chosen, the Language into Act Theory (Cresti 2000), which considers prosodic breaks in the transcription instead of orthographic punctuation marks, has shown to be fundamental to the analysis of hesitations and chunks, that is, groups of words frequently uttered

together, without pauses, continuously, as a string (Boers et al. 2006; Wood 2009). The investigation starts by verifying that native speakers use false starts, breaks and filled pauses as much as non-native speakers, which conforms to recent studies that even native speakers may produce "disfluencies" (McCarthy 2005). An apparent reason is the fact that silence may be confused with an offer to change the floor, when in fact the speaker wishes to hold it, contrary to studies which supply evidence that, in the flow of the interaction, hesitations seem to be a request for the interlocutor´s filling (McCarthy 2010). Such can be justified by restraints imposed by the radio handling.

The analysis then moves from examining the disfluencies to investigating the automation of language (Götz 2013) with a view to the production of chunks, on the basis that the more they are used, the better the fluency - and therefore proficiency - is perceived (Wood 2009). Some examples are 'let me know', 'I think', 'if you can', among others, giving prominence to pragmatic features. This paper concludes with the analysis of the turn-taking by examining initiating markers (Tao 2003), and with an observation of how pilots and air traffic controllers identify when to take the turn in radio communications.

Alderson, J. C. (2008). Final report on a survey of Aviation English tests. Retrieved from EALTA. Available at < http://www.ealta.eu.org/documents/archive/alderson_2008.pdf>.

Bachman, L. & Palmer, A. (2010). *Language assessment in practice*: Developing language assessments and justifying their use in the real world. Oxford: Oxford University Press.

Boers, F.; Eyckmans, J.; Kappel, J.; Stengers, H. & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a Lexical Approach to the test. *Language Teaching Research*, *10* (3), 245–261.

Cresti, E. (2000). *Corpus di italiano parlato*. v. 1. Firenze: Accademia della Crusca.

Emery, H. (2014). Developments in LSP testing 30 years on? The case of Aviation English. *Language Assessment Quarterly*, *11* (2), 198-215.

International Civil Aviation Organization. (2010). *Doc 9835*: Manual of Implementation of Language Proficiency. Montreal.

Götz, S. (2013). *Fluency in native and nonnative English speech*: Studies in Corpus Linguistics 53. Amsterdam: John Benjamins.

Lopez, S., Condamines, A., Josselin-Leray, A., O'Donoghue, M., & Salmon, R. (2013). Linguistic analysis of English phraseology and plain language in air-ground communication. *Journal of Air Transport Studies*, *4* (1), 44-60.

McCarthy, M. (2005). Fluency and confluence: What fluent speakers do. *The Language Teacher*, *29* (6), 26–28.

_____. (2010). Spoken fluency revisited. *English Profile Journal*, *1*(4).

Moder, C. & Halleck, G. (2009). Planes, politics and oral proficiency: Testing international air traffic controllers. *Australian Review of Applied Linguistics, 32* (3), 25.1-25.16.

Tao, H. (2003). Turn initiators in spoken English: A corpus-based approach to interaction and grammar. In: P. Leistyna & C. Meyer (Eds.), *Corpus Analysis: Language Structure and Language Use.* Amsterdam: Rodopi, 187–207.

Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics*, *12* (1), 39–57.

A qualitative analysis of perceptions of fluency in second language French

Yvonne PRÉFONTAINE and Judit KORMOS

Lancaster University, Department of Linguistics and English Language and Oracle Corporation

In the field of second language (L2) fluency, there is a common adherence to quantitative methods to examine characteristics and features of speech. This study extends the field by reporting on an investigation that analyzed native-speaker listeners' perceptions of L2 fluency in French from a qualitative perspective. Qualitative studies of L2 speech production and perception exploring French fluency are scarce within the SLA research literature. In fact, the only study to investigate L2 fluency in French from both a speech production and a perception perspective was conducted by Freed (2000).

The goal of our study was to compile and analyze three raters' qualitative perceptions of fluent performance produced by L2 learners of French on a set of oral narrative tasks. The research aimed to uncover the linguistic processing experience of listeners when they evaluate L2 speech. Our study sought to answer the question: What are the features of L2 learners' oral production that influence perceptions of L2 fluency in French? Our study sought to contribute to the existing literature by focusing on perceived fluency with regard to speech features of rhythm particularly, as well as speed, pauses, lexical retrieval, selfcorrection and efficiency/effortlessness.

The context for the study was a 5-week French immersion programme at a large francophone university in Québec, Canada. The L2 speaker participants were 40 volunteer undergraduate and graduate learners registered in beginning, intermediate and advanced French courses. There were approximately 13 participants per level, ranging in age from 18 to 69 years. The participants comprised 26 Canadian, 13 American and one British student, of whom 21 were female. The participants were all native speakers of English and varied in their exposure to French language study and the francophone world. They had an average of six years of French instruction in a regular classroom setting with the exception of 10 participants who had attended for an average of nine years in a French immersion setting in another Canadian province outside Québec.

The rater participants were three native speakers and French language instructors from the same university who were recruited to judge L2 speech production qualitatively. Although the raters had many years of experience in teaching French to non-francophone students, none had previously been involved in any L2 fluency rating projects.

The L2 speaker participants responded to three narrative speech tasks ranging in task complexity, demand and scope. In the first task, participants narrated a story based on six random pictures. The second task, a story retell, entailed retelling a story based on a short text in English about a horseback riding accident. In the final task, participants narrated a story based on an 11-frame cartoon strip presented in chronological order.

The three raters listened to each of 120 speech performances and gave their written qualitative impressions in which they described the features that most influenced their perceptions of L2 fluency in French. In the qualitative research conducted, the raters were intentionally untrained and were not provided with a definition of fluency to serve as a guide.

This procedure was implemented in order to avoid imposing a particular selffulfilling construct of L2 fluency on the raters. Rather, they were informed of the overall goal which was to reveal what native speakers "perceive in the real world as a listener" when they hear L2 spoken French and what personally influences their perceptions from both qualitative and quantitative perspectives. This open-ended approach allowed the raters to make their own judgments about what constitutes L2 fluency in French, while still providing considerable qualitative detail.

The speech-rating project was conducted online using Google Drive. Each participant's three speech productions were uploaded in a randomized order to ensure that the raters refrained from rating the same student equally across tasks. The raters provided written qualitative comments describing the fluency features that most influenced their evaluations of perceived fluency.

The data suggest that a fine balance must be struck between speed, pausing, lexical retrieval, self-correction, efficiency/effortlessness, and particularly rhythm, to qualify as a fluent L2 French speaker. However, as the excerpts from the qualitative comments indicate, these speech features and concepts are inherently intertwined and cannot easily be distinguished from each other. While numerous comments refer to speed of delivery, in general it is mentioned less often than rhythm, indicating that it might be secondary to rhythm in L2 French. According to the raters' reactions, an L2 speaker is considered fluent when they can combine all these features to speak easily, relatively quickly and with pauses at appropriate junctures. While all these factors together weave an intricate pattern in the fabric of L2 fluency, the speech features that were most frequently commented on by the raters in this dataset were speech rhythm, efficiency and effortlessness. This may be due to the fact that speech rhythms and efficiency/effortlessness are easily perceived features, and given the L2 French immersion context to which the raters are accustomed, they may be more conscious of these speech characteristics in L2 learners. Nonetheless, speech rhythm appears to be an overarching characteristic of fluent L2 speech in French.

Freed, B. 2000. Is fluency, like beauty, in the eyes and ears of the beholder? In H. Riggenbach (ed.), *Perspectives on fluency*. 243–265. Ann Arbor, MI: University of Michigan Press.

Is it pause type or pause length that matters? Hesitation fillers and their effect on fluency ratings and the perception of L2 speakers

Sandra REITBRECHT

University of Vienna, University College of Teacher Education Vienna

Introduction

In the context of second language learning, fluency is considered as an important feature of oral proficiency. Therefore, recent papers present tasks for a fluency training (e.g. Rossiter et al. 2010; Funk 2014) and prove their effect on fluency improvement (see de Jong/Perfetti 2011; Lambert et al. 2016). Furthermore, L2 fluency research focusses on the field of language testing and investigates the relationship between utterance and perceived fluency (e.g. Kormos/Dénes 2004; Rossiter 2009; Bosker et al. 2013; Préfontaine et al. 2015). Regarding the importance of

the number of silent and filled pauses for fluency ratings, divergent results are existent: On the one hand, Bosker's et al. study (2013, 167) shows that the number of silent pauses correlates better with the impression of disfluent speech than the number of filled pauses. On the other hand, Kormos/Dénes (2004, 155) report that in their data, the number of silent pauses and the number of filled pauses correlate both only weakly with the impression of disfluent speech. Referring to these results, this paper presents two case studies that investigate the effect of pause type (silent and filled pauses) (a) on fluency ratings and (b) on the speech-based perception of non-native speakers, their personality and their inner psychological state. This second research interest is positioned in the field of speech effect research (Sprechwirkungsforschung) which is based on the assumption of Bühler's Organon model that listeners can interpret the speech signal as symptoms (see Hirschfeld et al. 2010). Furthermore, recent L2 studies analysing verbal data also give evidence of the fact that listeners refer to L2 speakers' personality or their inner psychological state when rating accentedness or fluency (see Hayes-Harb/Hacking 2015; Préfontaine/Kormos 2016).

Study

To investigate the impact of the two different pause types (silent and filled pauses), the audiotaped performances of two learners of German as a foreign language on an oral narrative task were manipulated in the way that in the second version, the (majority of the) hesitation particles was/were replaced by silence and filled pauses thus became silent pauses (see table 1).

Table 1: Number of filled pauses per minute

	Original performance	**Manipulated performance**
Speaker / Case study 1	20,5	3,2
Speaker / Case study 2	21,0	0,0

In each case study, the original and the manipulated stimulus were presented to two different groups of listeners (native-speakers of German) who (1) answered an open-ended question concerning their impressions of the speakers and their performance and (2) rated fluency as well as certain speaker-related attributes (*assured*, *nervous*, *relaxed*, *well prepared*, *making an effort* and *competent*) in a six-point Likert scale (5 = *I totally agree*; 0 = *I totally disagree*).

Results

Table 2 shows the medians for *fluent* as well as for the above mentioned attributes. Grey fields indicate that the ratings for the original and the manipulated stimulus differed significantly (Mann-Whitney U test; $p < 0,01$). The medians show that the manipulation (replacing hesitation fillers as *äh*, *mmh*, *ähm* by silence) showed different effects in the two case studies: The L2 speaker in case study 1 was perceived as *more fluent*, *more assured*, *less nervous*, *more relaxed*, *better prepared*, *more competent and making more effort* when the raters listened to the original stimulus with filled pauses. In contrast, the medians for the L2 speaker in case study 2 differed only for three attributes, namely *nervous*, *relaxed* and *making an effort*. Furthermore, for *nervous*, *relaxed* and *making an effort*, the manipulation led to divergent changes in the two

case studies: Speaker 1 gained better results in the setting with the original stimulus, speaker 2, in contrast, in the setting with the manipulated data.

Table 2: Medians of raters' scores (5 = *I totally agree*; 0 = *I totally disagree)*

	fluent	*assured*	*nervous*	*relaxed*	*well prepared*	*competent*	*making an effort*
SPEAKER / CASE STUDY 1							
original (n=24)	3,0	3,0	2,5	2,0	4,0	3,5	4,5
manipulated (n=29)	1,0	1,0	3,0	1,0	2,0	1,0	4,0
SPEAKER / CASE STUDY 2							
original (n=26)	4,0	4,0	3,0	2,0	4,0	4,0	4,5
manipulated (n=26)	4,0	4,0	2,0	3,0	4,0	4,0	5,0

Discussion

One possible reason for the different/divergent impact of the change of pause type in the two case studies can be the concrete pause structures and the context in which the two L2 speakers produced the hesitation fillers. Whereas speaker 1 realized longer fillers and a considerable number of them in the context of silence (silence – hesitation filler – silence) or before silence (hesitation filler – silence), speaker 2 produced the majority of the fillers directly between words (word – hesitation filler – word). Thus, the mean duration of the silent pauses that were created by the manipulation differed strongly between the two case studies (1,3 seconds for speaker 1 and 0,4 seconds for speaker 2). Analysing the listeners' comments in the open-ended question also reveals that in case study 1, they strongly referred to the feature of pause length, less to the feature of pause type: Whereas for the original version, comments as *no long pauses* or *despite some reflexion pauses*, *no long interruptions* could be found, for the manipulated version, the listeners did not use any comparable positive comments and pause length was often mentioned in a negative way (e.g. *too long pauses* or *many break downs*).

Nevertheless, the quality of the manipulated data and their naturalness plausibility have to be mentioned as potential interfering aspects. Although the listeners had the possibility to comment on data quality at the end of the survey, none of them used this option. Thus, in future studies, a rating scale for naturalness plausibility has to be integrated into the questionnaire. Furthermore, it is interesting to take a closer look to the speech data itself: More than half of the silent pauses in the manipulated data in case study 1 reached a duration of more than 1,0 seconds and up to 4,7 seconds. The results of a larger corpus (Reitbrecht 2016) show that such durations are less frequent for silent pauses than for pause structures that include a hesitation filler. Qualitative aspects as breathing noise directly before a hesitation filler also have to be mentioned regarding the naturalness plausibility of the here used manipulated data. Furthermore, the number of listeners in this study was a limited one.

Although further studies are necessary to get better insights into the complex interaction of certain utterance fluency features in the perception of L2 speakers and the rating of L2 fluency, the results of these two case studies can be considered as an interesting contribution to the following two points: Firstly, they can be understood as an argument for a more detailed analysis of filled/combined pauses and their structures in the field of (L2) fluency research. Secondly, the results show clearly that in the context of second language teaching/learning, teachers should refrain from general or generalizing statements or advice regarding the use or avoidance of certain hesitation phenomena.

Bosker, Hans Rutger / Pinget, Anne-France / Quené, Hugo/Sanders, Ted / de Jong, Nivja H. (2013): What makes speech sound fluent? The contributions of pauses, speed and repairs. In: Language Testing 30/2, 159-175.

De Jong, Nel / Perfetti, Charles A. (2011): "Fluency Training in the ESL Classroom: An Experimental Study of Fluency Development and Proceduralization". In: Language Learning 61/2/2011, 533-568.

Funk, Hermann (2014): „Lernziel fremdsprachliche Flüssigkeit' – zur Problematik des Übungsdesigns im Fremdsprachenunterricht". In: Burwitz-Melzer, Eva/Königs, Frank G./Riemer, Claudia (Ed.): Perspektiven der Mündlichkeit. Arbeitspapiere der 34. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts. Tübingen: Narr Francke Attempto Verlag, 39-49.

Hayes-Harb, Rachel / Hacking, Jane (2015): Beyond rating data. What do listeners believe underlies their accentedness judgements? In: Journal of Second Language Pronunciation 1/1, 43-64.

Hirschfeld, Ursula / Neuber, Baldur / Stock, Eberhard (2010): Phonetische Sprechwirkungsforschung im Bereich der interkulturellen Kommunikation. In: Hirschfeld, Ursula/Stock, Eberhard (Ed.): Sprechwissenschaftlich-phonetische Untersuchungen zur interkulturellen Kommunikation Russisch-Deutsch. Peter Lang Verlag Frankfurt a. M. (Hallesche Schriften zur Sprechwissenschaft und Phonetik 33), 43-67.

Kormos, Judit / Dénes, Mariann (2004): Exploring measures and perceptions of fluency in the speech of second language learners. In: System 32, 145-164.

Lambert, Craig / Kormos, Judit / Minn, Danny (2016): Task Repetition and Second Language Speech Processing. In: Studies in Second Language Acquisition, online.

Prefontaine, Yvonne / Judit Kormos (2016): A qualitative analysis of perceptions of fluency in second language French. International Review of Applied Linguistics 54/2, 151-169.

Préfontaine, Yvonne / Kormos, Judit / Johnson, Daniel Ezra (2015): How do utterance measures predict raters' perceptions of fluency in French as a foreign language?. In: Language Testing (OnlineFirst), 1-21.

Reitbrecht, Sandra (2016): Häsitationsphänomene in der Fremdsprache Deutsch und ihre Bedeutung für die Sprechwirkung (PhD Thesis).

Rossiter, Marian J. (2009): Perceptions of L2 Fluency by Native and Non-native Speakers of English. In: The Canadian Modern Language Review 65/3, 395-412.

Rossiter, Marian J. / Derwing, Tracey M. / Manimtim, Linda G. / Thomson, Ron I. (2010): "Oral Fluency: The Neglected Component in the Communicative Language Classroom". In: The Canadian Modern Language Review 66/4/2010, 583-606.

Differences in second language speech fluency ratings: native versus nonnative listeners

Ralph L. ROSE

Waseda University Faculty of Science and Engineering

The perception of speech in cross-linguistic situations is known to be modulated by the listener's language background. In studies of the perception of foreign-accented speech by nonnative speakers (persons speaking in their nonnative language), nonnative listeners (persons listening to speech in their nonnative language) show greater comprehension than native listeners (persons listening to speech in their native language) (Bent and Bradlow 2003). Ratings of accentedness may also differ: Nonnative listeners judge foreign-accented speech less harshly than native listeners (Wester and Mayo 2014). Somewhat less studied is perceptions of fluency in similar cross-linguistic situations. Rossiter (2009) observed that fluency ratings by nonnative listeners of nonnative speech were lower overall than those given by native listeners. However, these rater groups were similar in that their ratings were most highly correlated with the speakers' articulation rate and silent pause frequency, suggesting that raters attend to these temporal parameters when rating fluency. In contrast, Foote and Trofimovich (2016) observed that native listeners depended more on silent pause frequency (over other factors) as a measure of fluency than did nonnative listeners, while the latter depended more on speech rate.

The diversity of language backgrounds in previous work could diminish effects associated with particular languages (cf., Bent and Bradlow's intelligibility benefit was larger for nativelanguage sharing participants). Therefore, the present study narrows the scope of nonnative listener to those listeners who also share the same native language as the speaker. This may allow effects pertaining to one specific language group to be revealed. While narrower in scope, the present study is an attempt to contribute to the wider body of evidence on native vs. nonnative fluency ratings.

The present work takes advantage of a crosslinguistic speech corpus (Rose 2013) in which native speakers of Japanese spoke for several minutes in response to various elicitation tasks in both their native language and English, their second language. These tasks included reading aloud and two types of spontaneous speech tasks: picture description—in which participants described a single frame scene or multi-frame sequence of scenes— and topic narrative—in which participants were given a topic (e.g., explain basketball to someone who has never seen it)..

For each speaker in the corpus (N=35 adults), seven thirty-second clips of their English (i.e., nonnative) speech were extracted from the three task types (three clips each from picture description and topic narrative and one clip from reading aloud). These clips were rated on a 9-point scale for fluency (1=low, 9=high) by native listeners (N=34 native English-speaking adults) using the Amazon Mechanical Turk crowd-sourcing work system. Raters were instructed to judge the "smoothness" of the speech and not other features like pronunciation or syntactic complexity. These results were previously reported in a study of the relationship between first and second language speech and fluency ratings (Rose 2015). In the present study, this earlier work was extended by having nonnative listeners (N=20 native Japanesespeaking adults) listen to the same set of recordings and rate the fluency in a similar manner.

The present study also analyzes two additional temporal parameters, filled pauses and repairs. The central comparison is fluency ratings by nonnative listeners (who share the same native language as the speakers) to those by native listeners.

Results show that the nonnative listeners gave lower fluency ratings overall (mean=4.4, sd=1.6) than did native listeners (mean=4.9, sd=1.7). This difference was significant according to a repeated-measures anova [$F(1,52)=7.6$, $p<0.01$]. This is consistent with previous findings and supports the view that nonnative listeners tend to judge nonnative speakers' fluency more harshly, although they share a native language.

In order to examine which features of speech that raters are attending to in their judgments, a step-wise linear regression was performed with fluency rating as the dependent variable. Independent variables included articulation rate (syllables per minute of phonation time), silent pause length (mean length of silent gaps of 300ms or greater; c.f., De Jong and Bosker 2013) , silent pause rate (number of silent gaps per minute), filled pause rate (number of occurrences of 'um' or 'uh' per minute), and repair rate (number of repair sequences per minute). The first three temporal parameters—articulation rate, silent pause length, and silent pause rate—were measured automatically using a Praat script (Boersma and Weenink 2013, Quené, Persoon, and De Jong 2011). The filled pause and repair rates were taken from transcriptions of the corpus. These variables plus elicitation task and rater language were included in the regression model.

The regression analysis shows that all factors were significant [$F(8,475)=118$, $p<0.001$; adjusted $R^2=0.66$] . Higher fluency rates are correlated with higher articulation rate, lower pause rate, shorter pause length, *higher* filled pause rate and lower repair rate. When separate analyses are performed on rater groups, articulation rate, silent pause length and filled pause rate remain for both groups while silent pause rate remains only for nonnative raters [$F(6,235)=104$, $p<0.001$; adjusted $R^2=0.72$] and repair rate remains only for native raters [$F(6,235)=58$, $p<0.001$; adjusted $R^2=0.58$].

The regression analysis further suggests differences between the reading aloud task and the other spontaneous speech tasks. After removing the reading aloud data, the overall models remain the same except that filled pause rate disappears from the optimal model for native raters [$F(3,203)=56$, $p<0.001$; adjusted $R^2=0.45$].

In contrast to previous findings, the present study suggests that nonnative raters judge fluency by somewhat different criteria than do native raters. While attention to articulation rate and pause length overlaps between them, nonnative raters also pay attention to pause rate while native raters pay attention to repair rate. This difference might be partially explained by differences in the perceptual tasks of the two raters. Previous research has shown that native listeners have a comprehension disadvantage (Bent and Bradlow 2003). Thus, more frequent pauses may actually aid their comprehension without being as intrusive as they are to nonnative listeners. On the other hand, once native listeners comprehend the speech, they are more likely to recognize repairs because it requires a higher level syntactic knowledge than some nonnative listeners have. Hence, the latter may simply be showing quicker word recognition, but lower syntactic comprehension than native raters.

The observed advantage for filled pause rate with nonnative raters may be related to crosslinguistic differences in filled pause use. Filled pauses have been shown to be generally more frequent in Japanese than English (Watanabe and Toyama 2016). Hence, a higher filled

pause rate may be regarded by these raters as somehow more natural, and therefore evidence of higher speaker fluency. But perhaps this becomes apparent to them only with the slower articulation rate of spontaneous speech.

Taken together with previous findings, the results here suggest that fluency ratings may be influenced by the native-language sharing status of listeners and raters. However, since the present study did not include non-native-language-sharing nonnative raters, a definitive conclusion cannot be made. This is the focus of future work.

Amazon Mechanical Turk web site. www.mturk.com

Bent, T. and Bradlow, AR. (2003) The interlanguage speech intelligibility benefit. In *Journal of the Acoustical Society of America*, 114(3): 1600-1610.

Boersma, P., Weenink, D. (2013): Praat: doing phonetics by computer [Computer program]. Version 5.4.03 , retrieved 26 Dec 2014 from www.praat.org

De Jong, N. H., & Bosker H. R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In *Proceedings of The 6th Workshop on Disfluency in Spontaneous Speech (DiSS 2013)*, p. 17-20.

Foote, JA. and Trofimovich, P. (2016) A Multidimensional Scaling Study of Native and Non-Native Listeners' Perception of Second Language Speech. In Perceptual and Motor Skills, 122(2): 470489.

Quené, H., Persoon, I., de Jong, N. Praat Script Syllable Nuclei v2 [Praat Script]. Version 28 Feb 2011, retrieved 26 Dec 2014 from https://sites.google.com/site/speechrate/Home/praatscript-syllablenuclei-v2

Rose, R. (2013) Crosslinguistic Corpus of Hesitation Phenomena: A corpus for investigating first and second language speech performance. In *Proceedings of the 14th Annual Conference of the International Speech Communication Association (Interspeech 2013)*, Lyon, France, pp. 992996.

Rose, R. (2015) Temporal Variables in First and Second Language Speech and Perception of Fluency. In *The Scottish Consortium for ICPhS 2015 (Ed.), In Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: the University of Glasgow.

Rossiter, MJ. (2009) Perceptions of L2 Fluency by Native and Non-native Speakers of English. In *Canadian Modern Language Review*, 65(3):395.

Watanabe, M. and Toyama, S. (2016) Comparison of filled pause durations and frequencies in parallel corpus of English and Japanese. *Programme of The 16th Australasian International Conference on Speech Science and Technology (SST2016)*, p.42. Parramatta, Australia.

Wester, M. and Mayo, C. (2014) Accent rating by native and non-native listeners. In Proceedings of International Conference on Acoustics, Speech, and Signal Processing, 7749-7753.

When and why do old speakers use more fillers than young speakers?

Lucie ROUSIER-VERCRUYSSEN[1,2], Anne LACHERET[2], Marion FOSSARD[1]

University of Neuchâtel[1], University of Paris Ouest Nanterre la Défense[2]

Introduction

Fillers (‘euh’, ‘hum’), which occur frequently in conversations, are produced in specific distributions within discourse (Campione & Véronis, 2005; Goldman et al., 2010). According to Arnold (2010), they indicate discourse or utterance planning difficulties. In a storytelling

task, speakers are more disfluent when they introduce new information (Fraundorf & Watson, 2014), especially when the storytelling is referentially more complex (1 vs. several characters) and ambiguous (same-sex characters, Arnold & Griffin, 2007; Arnold, 2010). Moreover, some studies found that older speakers produce more fillers than younger speakers (Bortfeld et al., 2001; Tottie, 2011; Laserna, Seih & Pennebaker, 2014). For these authors, the use of fillers can mark the speakers' age and signal a planning load. Since disfluencies are essentially related to the introduction of new information, the issues are the following: 1) are fillers more produced when referential complexity and/or ambiguity increases, i.e. is the difficulty to mention a referent in the presence of competitors expressed more by fillers than without? 2) are fillers linked to / influenced by executive functions such as planning abilities, i.e. is the use of fillers less present when executive functions are high?

The major aim of this study is to establish whether fillers produced in a storytelling task preferentially occur at strategic points of reference processing according to the accessibility of the referent, which would indicate an increased difficulty in discursive planning for younger (YS) and older speakers (OS).

Specifically, our goal is to (i) study the effect of referential complexity (1 vs. 2 characters) and referential ambiguity (2 characters of different sex vs. 2 characters of same sex) on the rate of fillers depending on the discourse stage (introduction, maintain and shift stage of characters), (ii) examine the relation between the rate of fillers and executive functions (inhibition, flexibility and planning) measured independently of the storytelling task, and (iii) compare the productions of fillers in YS and OS.

Method

The experiment, carried out with 30 YS (age span: 19-39) and 30 OS (age span: 59-79), is based on a storytelling in sequence task using the paradigm of referential communication (Clark & Wilkes-Gibbs, 1986). In this paradigm, the speaker has to tell a story based on a sequence of pictures so that the interlocutor sorts the pictures in the same order as the speaker. The experimental material is composed of sequences structured around 6 pictures: 2 sequences containing 1 character, 2 referentially more complex sequences containing 2 characters of different sex, and 2 referentially ambiguous sequences with 2 characters of the same sex. By manipulating the salience of characters (foreground or background within pictures) we could create 2 to 3 discourse stages: introduction, maintain and shift of characters. 360 storytellings, extracted from the SNF's data n°142069, a total duration of 8 hours and 30 minutes, were transcribed using Praat (Boersma & Weenink, 2009) and segmented into syllables with EasyAlign (Goldman, 2011). Each syllable is annotated as filler or not (Lacheret and al., in prep.). The rate of fillers is calculated as follows: the number of fillers *divided by* the total number of syllables, in each discourse stage of all the storytelling sequences. The participants were tested for the following cognitive abilities: inhibition (Stroop, 1935), flexibility (Kaplan et al., 1983) and planning (Wilson and al., 1996) and their scores were used to study the influence of these abilities on the rate of fillers. Five multiple linear regression analyses were conducted: one for the rate of fillers produced by YS, one for the rate of fillers produced by OS, one with cognitive measures for YS, one with cognitive measures for OS, and the last one to compare the production rate of fillers between YS and OS. For the first two, discourse stages were included as fixed factors. Referential complexity and referential ambiguity were also included as fixed factors in order to examine interactions. Two interactions were also included:

the first one between discourse stage and referential complexity, and the other one between discourse stage and referential ambiguity. For the second two models, executive functions were included as fixed factors. And for the last one, we included groups and discourse stages, determined as fixed factors, and the interaction between groups and discourse stages. Starting with the full model, we used model comparisons to determine whether the inclusion of a fixed factor and of an interaction was justified by the data. Only the final models will be presented.

Results

For each group, results show a significant effect of the introduction and shift stages compared to the maintain stage, with an increased rate of fillers (β=1.14, $p<0.05$ for YS, and β=1.97, $p<0.05$ for OS). This suggests that the rate of fillers is higher in the introduction and in the shift stages compared to the maintain stage. No significant interaction between discourse stages and referential complexity or ambiguity was found for YS and OS. It thus appears that the higher rate of fillers is not marked in a specific discourse stage when there is referential complexity or ambiguity. Furthermore, analyses conducted in both groups between the cognitive measures and the rate of fillers revealed that the more an OS is flexible, the less likely fillers are to be produced in the shift stage (β =-0.36, $p<0.05$). Moreover, the comparison between YS and OS indicates that OS produce significantly more fillers compared to YS (β=1.47, $p<0.05$), suggesting that all things being equal, OS are 4 times more likely to produce fillers compared to YS.

Conclusions

Our data reveal that fillers occur at strategic moments of reference processing: OS and YS produce more fillers when 1) they introduce a new character in discourse (Fraundorf & Watson, 2014), 2) they shift toward a less salient character (i.e. topic shift). Compared to the maintain stage, the introduction and shift stages generate a more important cognitive effort because the speaker has to mention a less accessible referent. Thus, the use of fillers indicates planning difficulties for these two discourse stages. Surprisingly, fillers produced by YS and OS do not increase in contexts of referential complexity or ambiguity for a given stage. So, speakers do not express their difficulty to mention a referent in the presence of competitors at different discourse stages by using fillers. Finally, our results highlight the influence of flexibility capacities on the rate of fillers for OS: OS with a lower flexibility use more fillers to express their difficulty in shifting characters. Their filler productions mark their difficulties to reorient the mention of a maintained referent towards another one. Further studies are needed to better characterize the relations between discourse planning, cognitive abilities and other types of disfluency.

Arnold, J., & Griffin, Z. 2007. The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language 56*(4): 521-536.

Arnold, J. E. 2010. How speakers refer: The role of accessibility. *Language and Linguistics Compass* 4: 187-203.

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & S. E. Brennan. 2001. Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech* 44 (2): 123-147.

Boersma, P., & D. Weenink, 2009. Praat: doing phonetics by computer (Version 5.1.05).

Campione, E., & J. Véronis. 2005. Pauses and hesitations in French spontaneous speech. *Proceedings of DiSS'05*: 43-46.

Clark, H. H., & D. Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition* 22: 1-39.

Fraundorf, S. H. & Watson, D. G. 2014. Alice's adventures in um-derland: psycholinguistic sources of variation in disfluency production. *Language, Cognition and Neuroscience 29*: 1083-1096.

Goldman, J.-P., Auchlin, A., & M. Avanzi. Submitted. Hesitations in read vs. spontaneous French in a multi-genre corpus. DiSS-LPSS Joint Workshop, Interspeech Satellite Event.

Goldman, J.-P. 2011. EasyAlign: an automatic phonetic alignment tool under Praat. *Proceedings of InterSpeech, September 2011, Firenze, Italy*.

Lacheret, A., Kahane, S., & P. Pietrandrea. Forthcoming. *Rhapsodie: A Prosodic and Syntactic Treebank for Spoken French*.

Laserna, C. M., Seih, Y.-T., & J. W. Pennebaker (2014). Um . . . Who Like Says You Know : Filler Word Use as a Function of Age, Gender, and Personality. *Journal of Language and Social Psychology.*

Kaplan, Goodglass & Weintraub. 1983. Boston Naming Test. *Philadelphia: Lea & Febiger*Stroop, J. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology.*

Tottie, G. (2011). Uh and um as sociolinguistic markers in British English. *International Journal of Corpus Linguistics*, *16*, 173-197.

Wilson, B., Alderman, N., Burgess, P., Emslie, H., & J. Evans. 1996. *Behavioural Assessment of the dysexecutive Syndrome (BADS)*. Thames Valley Test Company.

Disfluencies in Trump and Clinton first presidential debate

Ye TIAN

Université Paris-Diderot

The first presidential debate on 26 Sept 2016 between Donald Trump and Hillary Clinton was the most watched debate in American history. This paper analyse the disfluency patterns in this debate.

We annotated filled pauses, repetitions, self-repairs and abandoned utterances of the two speakers in the 90min detabe, discounting disfluencies during cross talk. It is an impression of many that Trump was more disfluent than Clinton. We found that in fact, the overall occurrences of disfluency are similar between the two, at 60 for Clinton and 73 for Trump. Taken into account the number of words uttered (Trump used 8866 words, while Clinton used 6580 words), the average rates of disfluency are almost the same: 7.6 per thousand words for Trump and 8.1 per thousand words for Clinton. Note that these rates are significantly lower than those in natural conversations in informal settings, estimated to be around 60 per thousand words by Fox Tree, 1995, or 12 per thousand words by (Bortfeld, Leon, Ì, Bloom, & Schober, 2001).

What WAS different is the types of disfluencies. Trump had a fair number of repetitions, repairs and abandoned utterances, and relatively few filled pauses. Clinton, on the other hand, had almost no repairs or abandoned utterances. She had a few repetitions, and many more filled pauses.

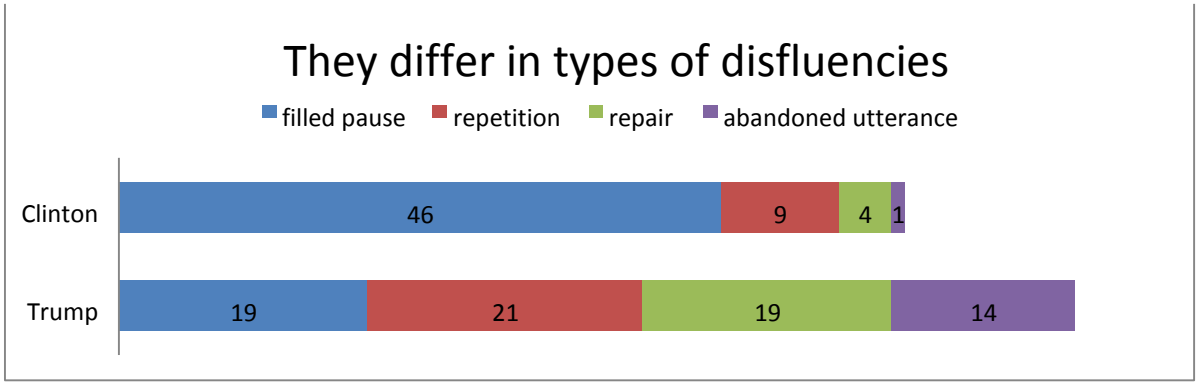


Figure 1 disfluency distribution of Trump and Clinton

Trump tends to repeat himself, he often stops mid-sentence to add something, and may or may not come back to his original partial sentence. The types of his disfluencies are perhaps why he doesn't come across as a well-prepared and eloquent speaker. Below are some examples:

Repetition:

TRUMP: New York -- New York has done an excellent job. And I give credit -- I give credit across the board going back two mayors.

Repair:

TRUMP: They (left + fired) 1,400 people.

TRUMP: I could name + { I mean} there are thousands of them.

Abandoned utterance:

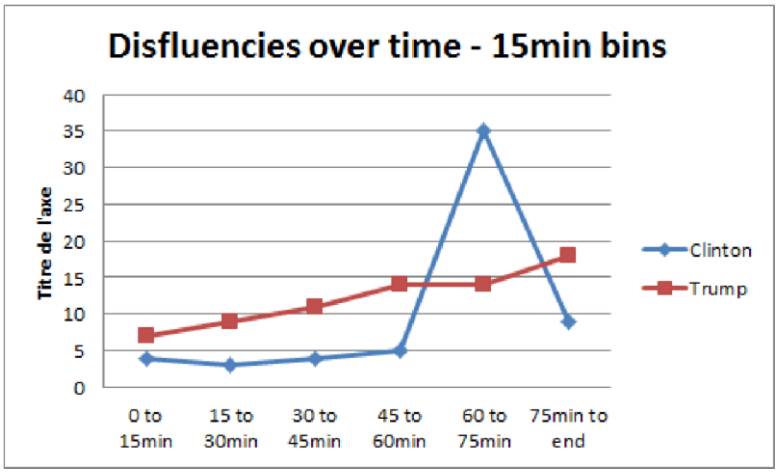
TRUMP: The African-American community -- because -- look, the community within the inner cities has been so badly treated.

TRUMP: whether it's -- I mean, I can just keep naming them all day long -- we need law and order in our country.

Clinton uses more filled pauses, but barely any other types of disfluencies. The filled pauses were not evenly distributed. There are pockets of utterances where filled pauses are more frequent. For example:

*CLINTON: Well, I think we need to do much more **{F uh}** with our tech companies to **{F uh}** prevent ISIS and their operatives **{F uh}** from being able to use the Internet to radicalize, even direct **{F uh}** people in our country and Europe and elsewhere. But we also have to intensify our air strikes against ISIS **{F uh}** and eventually support our Arab and Kurdish **{F uh}** partners to be able to actually take out ISIS **{F uh}** in Raqqa. **{F uh}** And we're hoping that **{F uh}** within the year we'll be able to push ISIS out of Iraq and then, you know, really squeeze them in Syria.*

Disfluency rates over time clearly shows that while Trump's disfluency rate increases steadily, Clinton's disfluency rate fluctuates, peaking between 60 to 75min, during which they were discussing fighting cyber-crime and ISIS.



From the perspective of perception, Trump style disfluency likely incurs more processing costs, for the listeners, and as a result affects the perception of discourse coherence. Levelt (1983) suggests that selfrepairs pose a "continuation problem" for the listeners, who must determine the start of the reparandum (the to-be-repaired), and how the repair replaces the reparandum. Similarly, upon listning to abandoned utterances, listeners may not know whether it was a reparandum and thus should be held in memory, or whether it should be discarded, and processing should start from scratch. These factors likely contributed to the perception that Trump was difficult to follow and incohenrent.

On the other hand, Clinton style disfluency, namely filled pauses only, may sometimes aid comprehension. Studies have shown that filled pauses may indicate discourse structure. Participants remember a story better after listening to "Alice in wonderland" with filled pauses than with coughs (Fraundorf & Watson, 2011), suggesting that compared to repairs and abandoned utterances, filled pauses have less negative impact on processing cost. However, (Brennan & Williams, 1995) showed that compared to silent pauses, filled pauses are perceived to indicate less confidence in the speaker. It is possible that Clinton's frequent filled pauses during discussion of cyber-crime and ISIS have lead to a perception of her low-confidence on this topic.

From the perspective of production, repetitions, repairs and abandoned utterances signal less initial planning and constant self-monitoring. They suggest that many of Trump's utterances were not prerehearsed, and his utterances reflects his thought processes – where often new and/or periphery ideas get activated and replace the original thought. On the other hand, filled pauses reflect planning of up-coming words/utterances. The lack of repairs and relatively higher rate of filled pauses of Clinton suggests that she has rehearsed some of the verses, and for those that were not well rehearsed, Clinton put in extra effort in planning. Overall, Trump's style is more egocentric, while Clinton's style is more listeneroriented.

In the final version of the paper, we aim to incorporate data from the second and last debates, as well as relating disfluency patterns to the content of their speech.

Bortfeld, H., Leon, Ê. S. D., Ì, S. E. B., Bloom, J. E., & Schober, Ë. M. F. (2001). D isfluency R ates in Conversation : Effects of Age , R elationship Topic , Role, and Gender. *Language and Speech*, *44*(3), 123–147.

Brennan, S., & Williams, M. (1995). The feeling of Another′ s Knowing: prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*.

Fraundorf, S., & Watson, D. (2011). The disfluent discourse: Effects of filled pauses on recall. *Journal of Memory and Language*.

LEVELT, W. (1983). Monitoring and self-repair in speech. *Cognition*.

Tree, J. F. (1995). The Effects of False Starts and Repetitions on the Processing of Subsequent Words in Spontaneous Speech. *Journal of Memory and Language*, *34*(6), 709–738.

The distribution of editing phrases in German, French and Chinese dialogues

Ye TIAN[1], Julian HOUGH[2], David SCHLANGEN[2] & Jonathan GINZBURG[1]

Université Paris-Diderot[1], Bielefeld University[2]

1. Introduction

In natural conversations, speakers frequently produce lexical and non-lexical filled pauses, both during hesitations, and within self-repairs. (Ginzburg et al., 2014) categorize disfluencies into ones that are *backwards looking* and ones that are *forwards looking*. Forward looking disfluencies are cases when an utterance is interrupted by a filled or silent pause, but are continued without an alteration. Backwards looking disfluencies are cases when an utterance is interrupted and replaced with an alteration that refers back to an already uttered reparandum, and an editing phrase (EP) is often inserted. we define EPs not by their lexical meaning, but by its structural context. Any ”words” used between a reparandum and its repair is considered an EP.

This paper introduces a multi-lingual natural dialogue corpus annotated for disfluency, and presents a preliminary results on the repertoire of filled pauses and EPs in three languages: French, Chinese and German.

2. Data and transcription

We use the DUEL corpus (Hough et al., 2016), consisting of 24 hours of natural, face-to-face, loosely task-directed dialogue in German, French and Mandarin Chinese. The corpus is uniquely positioned as a cross-linguistic, multimodal dialogue resource controlled for domain. DUEL includes audio, video and body tracking data and is transcribed and annotated for disfluency, laughter and exclamations. The data consists of 10 dyads per language. Transcription was done from the WAV audio files using Praat (Boersma and Weenink, 2010), following the instructions of the DUEL transcription and annotation manual (Hough et al., 2015), which specifies language general practices such as segmentation, disfluency annotation and laughter annotation, as well as language specific instructions regarding filled pauses, exclamations, and non-standard orthography.

2.1. Editing Phrase and Repair Annotation

Our annotations follow the light-weight inline method of dialogue annotation described by Hough et al. (2015). We utilize the disfluencies marked up as EPs (a class which includes filled pauses).

The filled pauses are annotated by a {F }, bracketing other fillers and editing terms simply with { } - e.g. I { you know } like her.

The inventory of EPs and filled pauses differ depending on the language. For example, in German, the common filled pauses are {F ah¨ }, {F ahm¨ } and {F hm}; in French they are {F euh}, {F mmh} and {F euhm}; in Chinese, they are {F en}, {F eh}, as well as demonstratives {F nage} (literally "that") and {F zhege} (literally "this"). For repairs, restarts and abandoned utterances, we mark the structure according to this scheme, consistent with the Switchboard repair mark-up (Meteer et al., 1995): (*reparandum* + { *EP* } *repair*)

3. The distribution of editing phrases across languages

3.1. Filled pauses

The distribution of filled pauses were summarized in tables 1,2 and 3. They are the most frequent in French, at 0.29 filled pauses per utterance, compared to 0.17 per utterance in Chinese and 0.13 per utterance in German. "Non-lexical" vowel based filled pauses such as "euh", "eh" and "ah"¨ are the most frequent filled pauses in all three languages. Certain "discourse markers" have similar distributions as those "non-lexical" filled pauses, e.g. "bah" (an interjection) in French, "ranhou" ("then") in Chinese, and "also" ("so") in German.

Table 1: French filled pauses

Filler	Occurrences	Percentage
euh	4089	60 %
bah	651	9 %
hein	291	4 %
genre	279	4 %
tu vois	260	4 %
Rmmh	255	4 %
ah	248	4 %
'fin	199	3 %
euhm	192	3 %
en fait	138	2 %
bon	135	2 %
ouais	128	2 %

Table 2: Chinese filled pauses

Filler	Occurrences	Percentage
eh	1066	28 %
ranhou (then)	578	15 %
en	541	14 %
jiushi (it is)	514	14 %
nage (that)	483	13 %
ah	304	8 %
em	284	7 %
zhege (this)	50	1 %

Table 3: German filled pauses

Filler	Ocurrences	Percentage
ah¨	698	48%
ahm¨	413	29%
also	124	9%
Hm/mh/uhm	56	4%
Fah/F¨ ahm¨	38	3 %
Oh	32	2 %
ach	15	1 %
achso	11	1%
ja	8	1%

3.2. Editing phrases

We extracted instances of disfluencies marked in the form of (reparandum + optional EP repair). They include repetitions and repairs. French and Chinese were similar in the rates of repetitions/ repairs. In French, 19% of utterances contain repetitions/repairs. There was a total of 3684 occurrences, on average 1.2 per utterance. In Chinese, 17% of utterances contain repetitions/ repairs. There was a total of 4476 occurrences, on average 1.15 per utterance. In contrast, only 8 %of utterances in German contain repetitions/repairs. There was a total of 1125 occurrences, on average 1.2 per utterance. In terms of EPs, French uses them more frequently than Chinese and German. 25% (French), 14% (Chinese) and 13% (German) of repetitions/ repairs used an EP. Both filled pauses and lexical items can be used as EPs. Few of the frequent lexical EPs contain in their meaning the sense of ”editing” or ”correction”. Tables 4, 5 and 6 summarize the distributions of EPs in three languages.

Table 4: French editing phrases

French	Occurrences	Percentage
euh	655	72%
ouais (yeah)	44	5%
genre (like)	42	5%
bah	29	3%
’fin (lastly)	23	3%
euhm	21	2 %
enfin (lastly)	16	2 %
voila	15	2 %
tu vois (you see)	13	1 %
bon (good)	8	1 %
donc (so)	8	1 %

Table 5: Chinese editing phrases

Editing phrase	Occurrences	Percentage
e	192	32 %
jiushi (is)	130	21 %
nage (that)	97	16 %
ranhou (then)	77	13 %
dui (correct)	30	5 %
em	22	4 %
non-verbal teeh noise ”tze”	16	3 %
oh	12	2 %
zhege (this)	11	2 %
shenme (what)	9	1 %
bushi/budui (no)	9	1 %

Table 6: German editing phrases

Editing phrase	Occurrences	Percentage
ah¨	82	56 %
ahm¨	30	20 %
also	19	13 %
ja	4	3 %

4. Conclusion and future work

We analyzed filled pauses and editing phrases in a multilinguistic dialogue corpus DUEL. We found that different languages use filled pauses and EPs are different rates (more frequent in French than in Chinese and German). A repertoire of both ”non-lexical” and lexical filled pauses were used in all three languages, and most of these filled pause can also function as EPs. For the final version of the paper, we aim to distinguish repetitions and different types of repairs (e.g. phonological, morphological, semantic, pragmatic), and investigate whether different EPs are preferred in different contexts.

Boersma, P. and Weenink, D. (2010). Praat: doing phonetics by computer.

Ginzburg, J., Fernandez,´ R., and Schlangen, D. (2014). Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics*, 7(9):64.

Hough, J., de Ruiter, L., Betz, S., and Schlangen, D. (2015). Disfluency and laughter annotation in a light-weight dialogue mark-up protocol. In *The 6th Workshop on Disfluency in Spontaneous Speech (DiSS)*.

Hough, J., Tian, Y., de Ruiter, L., Betz, S., Schlangen, D., and Ginzburg, J. (2016). DUEL: A Multi-lingual Multimodal Dialogue Corpus for Disfluency, Exclamations and Laughter. In *10th edition of the Language Resources and Evaluation Conference*.

Meteer, M. W., Taylor, A. A., MacIntyre, R., and Iyer, R. (1995). *Disfluency annotation stylebook for the switchboard corpus*. University of Pennsylvania.

Pauses in disfluent phases in native and non-native read speech

Jürgen TROUVAIN

Saarland University

Most studies on fluency and disfluency in second language (L2) research investigated non-scripted speech. This paper focus on read (or scripted) speech, more particularly on the less frequently studied topic of pauses in fluent and disfluent phases in native and non-native read speech.

Why study scripted speech? In contrast to unscripted speech in scripted speech important planning steps are missing or reduced, such as the conceptual preparation of the formulation, the morphosyntactic encoding, the selection of words, and aspects of social interactivity (Levelt 1989, Segalowitz 2010). Although scripted speech seems to be less demanding than non-scripted speech, reading aloud is used in various language registers. For instance, on the side of speech production lectures and talks are often fully or partially scripted, on the side of speech perception most broadcasted speech in television and radio is scripted, e.g. the news. It thus forms a part of everyday life.

A big advantage of read speech is that it provides a high level of comparability. It makes comparisons of various performances of the same text easier, e.g. between learners in a group, between samples of the same learner of two time points in the learning process, or between samples of the same speaker in her/his first language (L1) and second language (L2).

It is typical for spontaneous speech that a great number of disfluencies including and 'filled pauses', i.e. pauses containing filler particles like "uh" or "uhm", can be found, followed by repetitions and deletions (e.g. Shriberg 2001, Lickley 2015). In read speech, pauses at inappropriate locations (e.g. within a noun phrase), prolongations of single sounds, and particularly articulation errors (usually with repairing them, either with or without a short silent pause) are often found forms.

In comparison to native speech non-native speech usually shows a lower degree of utterance (or production) fluency. This is reflected by pausing, tempo and the use of disfluencies. L2 speakers usually produce more and longer pauses than L1 speakers. Together with a slower articulation rate this yields also in a slower speaking rate (which includes pauses).

Markers of the reduced utterance fluency in L2 speech can be found in non-scripted as well as scripted speech styles. One important difference between both speech modes seems to be the sometimes mentioned infrequent use of filler particles in read speech (e.g. Cucchiarini, Strik & Boves 2002).

For the present study we analysed a subset of the IFCASL bilingual phonetic learner corpus (Trouvain et al. 2016). 20 German and 20 French native speakers read the same narrative text ("The three little pigs") in their respective L1 and L2. Both language versions contain 13 sentences. In both language groups, 10 speakers were at beginner's level of L2 and 10 at advanced level (A2 vs. B2/C1, respectively, in the Common European Framework of Reference for Languages). In total, 80 recordings were annotated and analysed with respect to silent pauses, breath pauses ('unfilled' pauses with inhalation noise) and sections containing disfluencies (disfluent phases) or not (fluent phases). Recording durations of the text readings were between 50 and 200 seconds.

One of the main results is that in our read material virtually no hesitation particle occurred, i.e. all pauses were 'unfilled' pauses (in fluency research terminology, e.g. Lickley 2015). However, most pauses were breath pauses, i.e. they are phonetically filled with audible inhalation noises (making the term 'filled pause' questionable from a phonetic point of view). Contrary to this general pattern of pauses at appropriate prosodic-syntactic locations, the great majority of pauses in disfluent phases were non-breath pauses which were also shorter than breath pauses.

As expected beginners were generally more disfluent than advanced learners. Likewise speech of advanced learners showed more disfluencies than native speech (compared to the own native speech and also to L1 speakers). Native speech was not free of disfluencies, a remarkable number of speakers showed also (few) disfluencies in their L1. Native and non-native speakers also slightly differed with respect to prosodic phrasing. Generally, individual differences were rather large, with advanced L2 learners showing a low fluency level and beginners with a high degree of fluency.

A more detailed knowledge of individual patterns of disfluency can be very important for automatic (and human) fluency assessment. It would be a general benefit if we would know more about the aspects on which the perceived fluency are based in the speech production of various types of speech situations, e.g. to correlate the utterance fluency of read speech with that of spontaneous speech, as already done by Cucchiarini et al. (2010). Furthermore, it would be a benefit for an individual L2 learner to hear whether s/he has improved on utterance fluency and to use visualisations and other forms of feedback, e.g. in a computer-assisted pronunciation and fluency training.

Although in read speech the planning processes are basically reduced to the phonological planning and its phonetic execution, this type of data provides enough material to investigate L2 fluency. An interesting topic for future studies would be a comparison of L1 and L2 read and spontaneous speech of the same speaker.

Cucchiarini, C., Strik, H. & Boves, L. 2002. Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. Journal of the Acoustical Society of America 111(6), pp. 2862-2873.

Cucchiarini, C., van Doremalen, J. & Strik, H. 2010. Fluency in non-native read and spontaneous speech. Proceedings of DiSS-LPSS Joint Workshop 2010, Tokyo, Japan.

Levelt, W. 1989. Speaking. Cambridge, Massachusetts: MIT Press.

Lickley, R.J. 2015. Fluency and disfluency. In: Redford, M.A. (ed) The Handbook of Speech Production. Hoboken, NJ: John Wiley, pp. 445-474.

Segalowitz, N. 2010. Cognitive Bases of Second Language Fluency. London: Routledge.

Shriberg, E. 2001. To 'errrr' is human: ecology and acoustics of speech disfluencies. Journal of the International Phonetic Association 31, 153–169.

Trouvain, J., Bonneau, A., Colotte, V., Fauth, C., Fohr, D., Jouvet, D., Jügler, J., Laprie, Y., Mella, O., Möbius, B. & Zimmerer, F. 2016. The IFCASL corpus of French and German non-native and native read speech. Proceedings 9th Language Resources and Evaluation Conference (LREC), Portorož, pp. 13331338.

The cognitive basis of stuttering in bilinguals

Lize VAN DER LINDEN[1], Bernadette PIÉRART[1], Marie-Pierre DE PARTZ[1], Wouter DUYCK[2], Caroline MOERENHOUT[3], & Arnaud SZMALEC[1,2]

Université catholique de Louvain[1], Ghent University[2], BVBA Algemene Aanpak Stotteren[3]

The ability to speak and communicate with other people is an essential skill to succeed in this society, so we might assume that a disruption in this ability will have a dramatic impact on daily life.

Stuttering is a speech disorder affecting 1% of adults (Bloodstein & Bernstein, 2008) and is associated with involuntary sound repetitions, abnormal hesitations or pausing before speech and the prolongation of certain sounds, usually vowels and semivowels (Foundas et al., 2004). An influential model for explaining the cause of stuttering is the vicious circle hypothesis of Vasic and Wijnen (2005). Accordingly, speakers who stutter suffer from a malfunctioning of the monitoring mechanism of the speech production system. This hypothesis proposes that three attention parameters of the speech monitor are malfunctioning. First, people who stutter (PWS) invest a huge amount of attention resources in monitoring compared to people who do not stutter (PWNS). Second, the monitoring of PWS is rigid and maladaptive. PWS are continuously looking for all kinds of signs of realized or imminent disfluencies. Third, the speech monitor of PWS is too strict. PWS reject speech that PWNS would consider as normal. Evidence for this hypothesis comes from dual task experiments. In dual task experiments, participants are asked to carry out two (or more) tasks simultaneously, such as to simultaneously process visual and verbal material. It has been found that stuttering frequency decreased in PWS when the focus of attention was drawn away from speech production with a secondary task (Vasic & Wijnen, 2005). However, there is also research suggesting the opposite effect of dual task execution on stuttering frequency (e.g. Bosshardt, 2002), suggesting an alternative hypothesis that the speech of PWS is more sensitive to interference from other attention-demanding activities than that of PWNS and that performing a secondary task while speaking increases stuttering frequency.

In these times of globalization and cultural exchange, the number of bilinguals, or individuals who regularly use two languages, is constantly increasing. The prevalence of bilingualism is even estimated to exceed 50% of the world population (Grosjean, 2010). The influence of bilingualism on stuttering, however, remains largely unknown and research mainly focused on stuttering within one single language. In the current study, we aimed exploring whether and how bilingualism affects stuttering severity. It is widely accepted that for a bilingual who does not master his two languages to the same degree (moderate bilingual), speaking in a second language (L2) is more attentiondemanding than speaking in the first language (L1) due to the requirement to inhibit a constant coactivation of the dominant language (Green, 1998). If stuttering and bilingual language control tap on the same attentional system and if stuttering is due to a hypersensitive monitor that devotes too much attention to speech monitoring, stuttering frequency should be lower in L2 for moderate bilinguals. For those bilinguals, attention must be dedicated to the inhibition of the more dominant L1 in order to speak in the less proficient L2. This language control necessity might reduce hypersensitivity to the speech plan, because attention must be divided between the inhibition of the L1 as well as to the monitoring of the speech, hence creating a dual-task setting. Furthermore, stuttering frequency should be lower for bilinguals who are equally proficient in both languages

(advanced bilinguals) compared to moderate bilinguals because they have to inhibit both their L1 and L2 to the same extent. Therefore, they are constantly inhibiting one language in order to use the other and hence, which thus constantly requires an amount of the available attention capacity.

In the current study, we tested these hypotheses in a sample of 30 bilingual PWS. 15 bilinguals were moderate bilinguals who did not master their languages to the same degree and 15 bilinguals were highly proficient in both of their languages. They were required to perform a network description task in both their first (Dutch) and second (French or English) language, once as a single task, and once under dual-task conditions where attention resources were depleted through the introduction of a secondary tone discrimination task. During the network description task, participants were asked to describe in full sentences the route that was taken by a red dot across the drawings and lines on the network at the pace of the moving dot, which was calibrated at the normal speech rate of a native speaker in the tested language. A comparison of the stuttering frequency during the network description task between the two groups and between the two languages was done. First, there was no difference in overall stuttering frequency between our moderate and advanced bilingual groups, in any language. Interestingly, we observed a decrease of stuttering frequency under dual-task conditions within the L1 for both groups, while the dual-task condition had no impact on the stuttering frequency within the L2. Together, these data are in line with the Vicious Circle hypothesis, which assumes a hypersensitivity of speech monitoring in people who stutter. Our data show that reducing the available attentional resources for speech monitoring by a secondary, non-linguistic task that requires a lot of attention, reduces stuttering frequency. Our data further show that the requirement of language control in bilingualism, on the other hand, doesn't influence stuttering severity. We discuss the implications of these findings for theorizing about the interaction between stuttering and bilingualism.

Bloodstein, O., & Bernstein, R. (2008). *A handbook on stuttering*. Clifton Park: Delmar Learning.

Bosshardt, H.-G. (2002). Effects of concurrent cognitive processing on the fluency of word repetition: Comparison between persons who do and do not stutter. Journal of Fluency Disorders, 27 , 93–114.

Foundas, A., Bollich, A., J., F., Corey, D., Hurley, M., Lemen, L., & Heilman, K. (2004). Aberrant auditory processing and atypical planum temporale in developmental stuttering. Neurology, 63 , 1640–1646.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67-81.

Grosjean, F. (2010). *Bilingual: Life and Reality*. Cambridge: Harvard University Press.

Vasic, N., & Wijnen, F. (2005). Stuttering as a minitoring deficit. In R. J. Hartsuiker, Y. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech*.

A cross-linguistic comparison of filled pause frequency at clause boundaries in English and Japanese

Michiko WATANABE[1], Shohei TOYAMA[2]

National Institute for Japanese Language and Linguistics[1], The University of Tokyo[2]

Introduction

Filled pauses (FPs) are believed to be relevant to online speech planning. When speakers need extra time for preparing upcoming speech, they are likely to use FPs. Speakers may utter FPs because they need more time to come up with suitable expressions or to gain time for conceptual planning. Swerts found that FP rate was higher at major discourse boundaries than at shallower boundaries [1]. His findings are compatible with the idea that FPs are more frequent when heavier conceptual planning takes place. This leads to the hypothesis that FP rate is higher at deeper boundaries. Holms observed higher FP rate at sentence boundaries than at clause boundaries in English and French [2]. The results support the hypothesis. On the other hand, research on Japanese FPs indicates that FP rate at sentence boundaries is not higher than the rate at strong clause boundaries, which does not support the hypothesis [3]. To confirm the discrepancy between these findings, we examined FP rates at sentence and clause boundaries in similar type of speeches in English and Japanese. If the locations with higher FP rates differ, it is possible to conjecture that FPs are used differently depending on the language.

Method

2.1. Corpus

The Japanese data were obtained from a subset of "The Corpus of Spontaneous Japanese (CSJ)" [4]. We sampled 10 male and 10 female speakers in their twenties and early thirties. They were instructed to talk about given topics such as "the happiest memory in my life" for 10 to 15 minutes. FPs such as "anoo" and "eeto" were labeled. Sentence boundaries and clause boundaries were marked. In Japanese, subordinate clauses always precede the main clause and connective particles are located clause-finally. The degree of dependency of subordinate clauses on the main clause is indicated by the type of connective particles. For example, subordinate clauses with certain type of connectives can have their own topics and/or subjects, while clauses with other type of connectives cannot. Boundaries after subordinate clauses which are more dependent on the main clause are called weak clause boundaries, and boundaries after subordinate clauses which are less dependent on the main clause are called strong clause boundaries, hereafter.

English speeches were also collected so that a contrastive study would be possible. The participants were 10 male and 10 female speakers of American English in their twenties and early thirties. The participants were instructed to talk about a given topic "the most memorable event in my life" for at least 10 minutes in front of a small audience. FP labels were given to "um" and "uh" in the transcription. The beginning and the end of main clauses, coordinate clauses, and adverbial clauses with finite verbs were marked. Sentence boundaries were marked additionally to clause boundaries by one labeler, when the labeler found boundaries strong based on the content of clauses and the prosody.

2.2. Procedure

We counted the total number of clauses and the number of clauses with clause-initial FPs immediately after the three boundary types for each speech and averaged over 20 speeches for each language. The three boundary types are as follows:

Japanese: 1) sentence boundaries; 2) strong clause boundaries; 3) weak clause boundaries

English: 1) sentence boundaries; 2) clause boundaries immediately after coordinate clauses, called coordinate clause boundaries hereafter; 3) clause boundaries immediately after adverbial clauses with finite verbs, called adverbial clause boundaries hereafter.

English boundary grouping is based on the preceding clause type to be concordant with Japanese boundary grouping.

Results

In English, clause-initial FP rate was highest at sentence boundaries (29 out of 127 sentence boundaries on average per speech, 24%), lowest at coordinate clause boundaries (4 out of 67 coordinate clause boundaries, 8%), and inbetween at adverbial clause boundaries (2 out of 11 adverbial clause boundaries, 17%). Only the rate at coordinate clause boundaries was significantly lower than the other two rates.

In Japanese, clause-initial FP rate was highest at strong clause boundaries (16 out of 48 strong clause boundaries on average per speech, 35%), which was significantly higher than the rate at sentence boundaries (13 out of 47 sentence boundaries, 26%) and the rate at weak clause boundaries (24 out of 82 weak boundaries, 29%). There was no significant difference between the latter two rates.

Discussion

In English, the rate of clauses with clause-initial FPs was highest at sentence boundaries, which supports the hypothesis that FP rate is higher at deeper boundaries. FPs seem relevant to speakers' cognitive load of conceptual planning in English. We presumed that adverbial clause boundaries are shallower than coordinate clause boundaries because the content of the main clause is already planned, at least to certain extent, at the beginning of the sentence and not much planning is needed at adverbial clause boundaries. From this presumption FP rate at coordinate clause boundaries should be higher than the rate at adverbial clause boundaries, but it was not. We should reconsider this presumption. FP rate at coordinate clause boundaries was lower than the rate at adverbial clause boundaries possibly because of conjunctions at coordinate clause boundaries. Speakers can gain time for preparing upcoming speech while they are saying "and", for example. On the other hand, because conjunctions are not used at adverbial clause boundaries, speakers are likely to need other items such as FPs.

FP rate was not highest at sentence boundaries in Japanese, which does not support the hypothesis. Instead, FP rate was highest at strong clause boundaries. It is possible that Japanese FPs are not so relevant to conceptual planning as English FPs are. Silent pauses (SPs) tend to be longer at deeper boundaries. Some part of long SPs, instead of FPs, may be being used for speech planning at sentence boundaries in Japanese. The rate at strong clause boundaries was significantly higher than the rate at weak clause boundaries, indicating that conceptual planning

is not irrelevant to FP occurrence. The results suggest that the type of problems that FPs are most sensitive to may differ depending on the language.

[1] Swerts M. "Filled pauses as markers of discourse structure", *Journal of Pragmatics* 30:485-496, 1998.

[2] Holmes, V. M. "A crosslinguistic comparison of the production of utterances in discourse", *Cognition* 54(2):169-207, 1995.

[3] Watanabe, M. *Features and Roles of Filled Pauses in Speech Communication -A corpus-based study of spontaneous speech*. Tokyo: Hituzi Syobo Publishing, 2009.

[4] The National Institute for Japanese Language and Linguistics. http://pj.ninjal.ac.jp/corpus_center/csj/misc/preliminary/index_e.html, accessed on 24 June 2016.

A new perspective in analysing task effects on fluency development in L2. Mandarin beyond the classroom

Clare WRIGHT

University of Leeds

The study reported here, examining longitudinal changes in L2 Mandarin fluency during Study Abroad (SA), thus aims to add to our understanding of fluency at the cognitive/utterance interface by reporting on data, gathered across four monologic and dialogic tasks, involving prepared and unprepared speech. Unusually, we used both CLAN and PRAAT methodologies to track lexical and grammatical development as well as highlight specific patterns in (dis)fluency indicators. Given the nature of Mandarin, calculations usually based on word or syllable are all here based on single characters. The combined methods allowed us reliably to compare changes in speed, breakdown and complexity of output (Skehan, 2014) – namely articulation rate, number and location of silent pauses, number and location of filled pauses, repairs, mean number of characters per utterance and lexical diversity. Participants were ten adult L1 English learners of Mandarin, assessed before and after a year's Study Abroad (SA). We found that across tasks, measures generally showed clear improvement: in particular mean number of characters per utterance, articulation rate, use of filled pauses improved significantly ($p < .01$). However, other measures such as repairs, mean length of silent pauses, location of silent pauses, length of utterance (MLU) and lexical diversity, differed between tasks, and were very variable across the group. In particular we found that significant between-task differences in the monologic tasks on most variables disappeared, but not for MLU or clause-internal pausing. In the dialogic tasks, we found consistently significantly shorter MLU at both times on both tasks compared to the monologic tasks ($p < .05$), that the prepared task had significantly slower articulation rate ($p < .05$) and shorter MLU (though *ns*) than the unprepared task, though with greater SD, and inconsistent improvement on pausing, confounding our predictions that preparation would aid both monologic and dialogic tasks, and that time spent abroad during SA would favour greater development in unprepared vs. prepared tasks.

This study highlights useful task effects when considering development of L2 fluency; more broadly it emphasises the value of collecting detailed longitudinal SLA-motivated datasets of language learners moving between instructed and immersed contexts (Du, 2013), and using

theoreticallygrounded comparisons of different task demands on cognitive and utterance fluency development. However, we also note the current lack of standardised reliable measures differentiating between task-based cognitive load, planning effects, and monologic vs. dialogic interaction (Pallotti, 2009; Tavakoli, 2016). We will address how further research could address some of the theoretical and methodological issues involved in analysing systematic linguistic development in L2 Mandarin, and their pedagogical implications.

Du, H. (2013). The development of Chinese fluency during study abroad in China. *The Modern Language Journal* 97(1): 131-143.

Freed, B., Segalowitz, N. & Dewey, D. (2004). Context of learning and second language fluency in French. *Studies in Second Language Acquisition* 26 (2): 275-301.

Kormos, J. (2011). Speech production and the Cognition Hypothesis. In Robinson, P. (ed.).

Second Language Task Complexity, 39-60. Amsterdam: John Benjamins.

Pallotti, G. (2009). Complexity, Accuracy, Fluency: Defining, refining and differentiating constructs. *Applied Linguistics* 30: 590–601.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge. Skehan, P. (2014). Limited attentional capacity, second language performance, and taskbased pedagogy. In Peter Skehan (ed.), *Processing perspectives on task performance*, 211–260. Amsterdam: John Benjamins.

Tavakoli, P. (2016) Fluency in monologic and dialogic task performance: challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching* 54 (2): 133-150.

Wright, C. and Zhang C. (2014). Examining the effects of Study Abroad on L2 Chinese development among UK university learners. *Newcastle and Northumbria Working Papers in Linguistics* 20: 67-83.

Predicting ESL learners' oral proficiency by measuring the collocations in their spontaneous speech

Jing XU

Cambridge English Language Assessment

1. Introduction

> [Words] are interconnected, not isolates … meaning is derived from context, and … collocation is key (Moon, 2008, p. 243).

Moon's words aptly describe the scaffolding role of collocation in a language. Many researchers have argued for the centrality of collocation to speech communication based on the assumption that effective use of prefabricated units saves a speaker's limited attentional resources and increases speech accuracy and idiomaticity (Bygate, 1987; Kormos, 2006; Schmitt, 2010; Wray, 2002). However, very limited research has explored ways to measure the collocation occurrences in learner speech and the predictive power of these measures for perceived oral proficiency. The current study is novel in that it measures three dimensions of L2 collocation production in spontaneous speech: accuracy, complexity, and fluency.

2. Research Questions

The goal of this study was to investigate the empirical performance of a series of theorybased collocation measures for predicting L2 oral proficiency. It posed two research questions. First,

what collocation measures can effectively differentiate among L2 speakers of different proficiency levels? Second, to what extent can collocation measurement predict expert judgment of oral proficiency?

3. Methods

The study focused on ten syntactic patterns of lexical collocations in spontaneous L2 English speech. They were adjective + noun, adverb + adjective, adverb + verb, noun + noun, noun of noun, noun + verb, verb + noun, phrasal verb + adverb, noun + phrasal verb, and phrasal verb + noun. The spoken data were obtained from SPEAK/TEACH, a college-level oral English test developed and used by Iowa State University. SPEAK assesses a candidate's ability to engage in everyday conversation whereas TEACH assesses a candidate's ability to disseminate discipline-specific academic information.

The study participants were sixty graduate students whose native language was Mandarin Chinese. They were drawn from four oral proficiency groups using a stratified sampling technique. Their oral responses in the SPEAK/TEACH test were marked on a scale of 0 to 300 by a panel of three certified examiners based on criteria covering pronunciation, vocabulary, fluency, and content. The inter-rater reliabilities were .89 for SPEAK and .91 for TEACH.

The collocations identified in learner speech were coded in three dimensions: accuracy, complexity, and fluency. The accuracy dimension included two subcategories: semantic accuracy (i.e., meaningfulness of the co-occurrence of root morphemes) and grammatical accuracy (i.e., morpho-syntactic error). The complexity dimension also contained two subcategories: transparency and restrictedness (i.e., substitutability of collocates). The fluency dimension had a single subcategory: automaticity (i.e., smoothness of oral collocation production).

Collocation coding was performed by nineteen faculty members and students in an Applied Linguistics program. Among them, fourteen were native English speakers and five were nonnatives. Non-native coders did not perform coding of semantic accuracy and restrictedness which, according to Schmitt (1998), requires native-speaker intuitions.

To answer the first research question, a one-way between-subjects analysis of variance (ANOVA) was performed on each collocation measure as a function of oral proficiency levels. Post hoc tests were also performed to further examine the differences. To answer the second research question, multiple regression was performed using collocation measures as the independent variables and human criterion scores of oral performance as the dependent variable.

4. Results

A significant effect of proficiency levels was found for use of acceptable collocations ($F(3, 56) = 10.276$, $p < .01$, $\eta^2 = .355$), highly restricted collocations ($F(3, 56) = 6.890$, $p < .01$, $\eta^2 = .270$), and choppy collocations ($F(3, 56) = 41.230$, $p < .01$, $\eta^2 = .688$). However, no difference was found in use of partially figurative collocations ($F(3, 56) = 2.704$, $p = .054$, $\eta^2 = .102$). In fact, the participants rarely used this type of collocations regardless of oral proficiency.

The variation in collocation measures across proficiency levels met theoretical expectations. Post hoc analysis indicated that proficient English speakers (Levels 1 and 2) generally used acceptable collocations more frequently than non-proficient speakers (Levels 3 and 4). The former also produced choppy collocation units more frequently than the latter while speaking (see Figure 1).

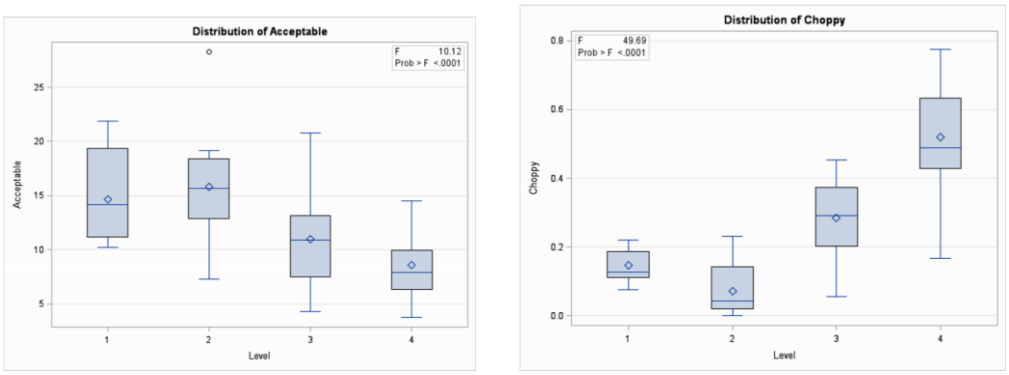


Figure 1. Box and whiskers charts showing score distributions of 'acceptable collocations' and 'choppy collocations'

When collocation measures were used to predict the sixty speakers' oral test scores, a significant prediction was found both in the SPEAK test ($F(8, 56) = 15.738$, $p < .001$, adjusted $R^2 = .678$) and the TEACH test ($F(8, 58) = 10.726$, $p < .001$, adjusted $R^2 = .573$). Particularly, there was a significant negative prediction of the SPEAK score by choppy collocations ($B = -74.69$, $t(56) = -6.07$, $p < .001$, $sr^2 = -.659$) and a significant negative prediction of the TEACH score by choppy collocations ($B = -84.90$, $t(58) = -6.410$, $p < .01$, $sr^2 = -.672$) and unacceptable collocations ($B = -6.70$, $t(58) = -2.446$, $p < .01$, $sr^2 = -.327$).

5. Conclusion and Implications

This study found that L2 speakers' collocation use in spontaneous speech had an effect on expert raters' judgment of their oral proficiency. The discussion over the collocationspeaking relationship has largely remained at a theoretical level (Millar, 2011). The findings of this study lent empirical support for the centrality of collocation competence to L2 oral proficiency.

In the common practices of L2 speaking assessment, a test taker's collocational performance is usually not specifically rated mainly because human raters can only focus on a limited range of speech characteristics. However, automated scoring technologies have given us high hope for looking into learner speech in great detail. This study suggests that the collocation occurrences in free L2 speech deserve computational linguists' closer attention because they contain useful information for predicting human judgment of oral proficiency.

Since construct underrepresentation is a major drawback of automated speaking assessment (Xi, 2010), it is recommended that computational linguists implement a multidimensional scale of collocation measurement to future development of speech auto-markers. Scoring L2 learners' collocation production in speaking practices and providing them with formative feedback on this performance would have a positive impact on language teaching and learning.

Bygate, M. (1987). *Speaking*. Oxford: Oxford University Press.

Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Millar, N. (2011). The processing of malformed formulaic language. *Applied Linguistics, 32*(2), 129-148.

Moon, R. (2008). Sinclair, phraseology, and lexicography. *International Journal of Lexicography, 21*(3), 243-254.

Schmitt, N. (1998). Measuring collocational knowledge: Key issues and an experimental assessment procedure. *I.T.L. Review of Applied Linguistics, 119-120*, 27-47.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. New York, NY: Palgrave Macmillan.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Xi, X. (2010). Automated scoring and feedback systems: Where are we and where are we heading? *Language Testing, 27*(3), 291-300.