Monday
May 15th, 2017

09,30	Opening	
10,00–10,45	**Plenary:** Valerie Hazan	*Spontaneous speech adaptations in challenging communicative conditions across the lifespan*
10,45–11,00	Questions	
11,00–11,30	Coffee break	
11,30–11,50	Vered Silber-Varod and Noam Amir	*"When two giants meet": The interaction between lexical stress and utterance-final prosody in spoken Hebrew*
11,50–12,10	Marko Liker	*Electropalatographic analysis of vowels in quasi-spontaneous speech*
12,10–12,30	Outi Tuomainen and Valerie Hazan	*Disfluencies in spontaneous speech in younger and older adults in easy and difficult communicative situations*
12,30–13,00	Questions	
13,00–14,00	LUNCH	
14,00–14,20	Alexandra Markó, Andrea Deme, Márton Bartók, Gergely Varjasi, Tekla Etelka Gráczi and Tamás Gábor Csapó	*Word-initial glottalization in the function of speech rate and vowel quality*
14,20–14,40	Sylvia Moosmüller, Hannah Leykum and Julia Brandstätter	*Is there a tendency to merge /e/ and /ɛ/ in Standard Austrian German? Data from read and spontaneous speech*
14,40–15,00	Krisztina Zajdó	*Building speech sounds through scaffolding: The case of motherese*
15,00–15,30	Questions	
15,30–16,00	Coffee break	
16,00–16,20	György Szaszák and Anna Moró	*Automatic punctuation recovery in read and spontaneous Hungarian using a recurrent neural network based sequential model for phonological phrases*
16,20–16,40	Judit Bóna	*Non-verbal vocalizations in spontaneous speech: The effect of age*
16,40–17,00	Dorottya Gyarmathy, Tilda Neuberger and Anita Auszmann	*The relationship between silent pause and breath-taking in spontaneous speech*
17,00–17,30	Questions	
19,00	**Welcome Party**	

Tuesday
May 16th, 2017

09,00–09,45	**Plenary:** Nick Campbell	*Towards interactive speech synthesis; an example of robot-human dialogues in a spontaneous environment*
09,45–10,00	Questions	
10,00–10,20	Katalin Mády, Uwe D. Reichel, Beáta Gyuris and Hans–Martin Gärtner	*The impact of syntax and pragmatics on the prosody of dialogue acts*
10,20–10,40	Tekla Etelka Gráczi, Alexandra Markó and Karolina Takács	*Word-initial glottalization in the function of articulation rate and word class*
10,40–11,00	Questions	
11,00–11,30	Coffee break	
11,30–11,50	Alexandra Bückins, Reinhold Greisbach and Anne Hermes	*Larynx movement in the production of Georgian ejective sounds*
11,50–12,10	Mária Gósy and Valéria Krepsz	*Phrase-final lengthening of phonemically short and long vowels in Hungarian spontaneous speech across ages*
12,10–12,30	Attila Gróf, Annamária Kovács, Anna Moró, Miklós Gábriel Tulics and Máté Ákos Tündik	*MagmaNet: Ensemble of 1D convolutional deep neural networks for speaker recognition in Hungarian*
12,30–13,00	Questions	
13,00–14,00	LUNCH	
14,00–14,45	**Plenary:** Ruth Huntley Bahr	*Variability in speech sound production: Covert contrasts in the speech of children with cochlear implants*
14,45–15,00	Questions	
15,00–15,20	Sarah Brandstetter	*Can you speak less dialect, please?*
15,20–15,40	Judit Bóna and Tímea Vakula	*Phonetic characteristics of disfluent word-repetitions: The effect of age and speech task*
15,40–16,00	Questions	
16,00–16,30	Coffee break	
16,30–16,50	György Szaszák and András Beke	*Exploiting prosodic and word embedding based features for automatic summarization of highly spontaneous Hungarian speech*
16,50–17,10	Davor Trošelj	*Vowel-formant frequencies of Hungarian–Croatian bilinguals and Hungarian monolinguals in spontaneous speech*
17,10–17,30	Questions	

Wednesday
May 17th, 2017

09,00–09,45	**Plenary:** Vesna Mildner	*Neurolinguistic aspects of speech processing*
09,45–10,00	Questions	
10,00–10,20	Tilda Neuberger and András Beke	*Effects of gemination on the duration and formant frequencies of adjacent vowels in Hungarian voiceless stops*
10,20–10,40	Ákos Gocsál	*Speaker age estimation by musicians and non-musicians*
10,40–11,00	Questions	
11,00–11,30	Coffee break	
11,30–11,50	Gordana Varosanec–Skaric, Zdravka Biocina and Gabrijela Kisicek	*Comparison of F0 measures for male speakers of Croatian, Serbian and Slovenian*
11,50–12,10	László Hunyadi	*On some linguistic properties of spoken Hungarian based on the HuComTech corpus*
12,10–12,30	Valéria Krepsz and Mária Gósy	*Stem and suffix durations in words of increasing length in children's spontaneous utterances*
12,30–13,00	Questions	
13,00–14,00	LUNCH	
14,00–14,20	István Szekrényes	*Challenges in automatic annotation and the perception of prosody in spontaneous speech*
14,20–14,40	Anita Auszmann	*A perceptual comparison: spontaneous speech of speakers' today and 40 years ago*
14,40–15,00	Mária Laczkó	*The temporal characteristics of teenagers in the various spontaneous speech genres*
15,00–15,30	Questions	
15,30–16,00	Coffee break	
16,00–16,30	Demonstration: Tamás Gábor Csapó, Andrea Deme, Tekla Etelka Gráczi, Alexandra Markó and Gergely Varjasi	*Synchronized speech, tongue ultrasound and lip movement video recordings with the "Micro" system*

17,00 **Farewell Cake and Drinks Party**

Abstracts

Monday

May 15th, 2017

Spontaneous speech adaptations in challenging communicative conditions across the lifespan

Plenary speech by

Valerie Hazan and Outi Tuomainen
Dept of Speech, Hearing and Phonetic Sciences, UCL, UK

Most of our knowledge about the acoustic-phonetic characteristics of speech come from speech production studies that have analysed controlled materials such as read sentences produced in isolation in a quiet environment. In typical communicative situations, the speech that we produce is likely to differ from such norms: it will be spontaneous, produced with true communicative intent, in less than ideal acoustic environments and quite often in a multi-tasking situation. In such situations, speech can be highly dynamic as we ongoingly adapt the level of clarity of our speech according to the demands of the communicative conditions, as suggested by Lindblom in his Hyper-Hypo model of speech production (Lindblom, 1990).

In our work, our aim has been to analyse the type of acoustic-phonetic adaptations made by speakers to counter the effects of adverse environments such as occur when communicating in noise, in the presence of other voices or with a hearing loss. We have recorded speech in laboratory conditions but have modelled natural communication by using a problem-based task that is carried out between two speakers. This picture-based 'spot the difference' task, called diapix (van Engen et al., 2010) involves the pair of speakers having to find 12 differences between their two pictures without seeing their conversational partner's picture. The degree of ease or difficulty with which speakers can communicate can be controlled by adding a communication barrier (e.g. babble noise, simulated hearing loss) affecting one or both of the speakers while they carry out the task together. This leads the person leading the interaction to naturally make adaptations to their spontaneous speech, producing a 'clear speaking style' in order to maintain effective communication, just as would happen in natural interactions. As speakers are recorded individually in connected sound-treated booths and communicate via headphones, a 'clean' and high-quality speech signal is recorded for each speaker.

In three consecutive large-scale studies, we have investigated such speech adaptations in 40 young adults (Hazan and Baker, 2011), 96 children aged 9 to 15 years (e.g., Hazan et al., 2016) and now 57 older adults aged 65 to 85 years, with 26 further younger adult controls (Tuomainen et al., 2016). A linked study has examined adaptations in hearing-impaired children while communicating with their hearing and hearing-impaired peers (Granlund et al., 2015). Each of these projects has led to the creation of large speech corpora (LUCID, kidLUCID and the forthcoming elderLUCID) containing many hours of spontaneous speech interactions. The lengthy processing of these corpora involves manual or automatic orthographic transcription, automatic alignment, manual checking of these alignments and the use of Praat scripts to obtain acoustic-phonetic measures. These measures include suprasegmental measures of articulation rate, fundamental frequency mean and range, relative intensity (representing spectral tilt) and segmental measures of vowel space.

In this talk, I will review our findings across these three studies spanning a broad age range; I will also discuss the challenges involved in the analysis of large spontaneous speech corpora. Our study with young adults showed that the adaptations that individual speakers made were, to an extent, dependent on the type of interference that was affecting their interlocutor (babble noise or vocoded speech), even though the speakers that we were analysing were not directly hearing the interference. This suggests that speakers used the direct or indirect feedback from their interlocutors during the interaction to attune their adaptations. Our study with children showed that they too made adaptations under similar conditions, although they had a tendency to use a strategy of increasing vocal effort (as shown by strong correlations between increases in fundamental frequency and decreases in spectral tilt) rather than using more varied strategies favoured by adults. Our ongoing study with older adults is showing a similar trend: older adults with age-related hearing loss tended to increase vocal effort to counter the effects of adverse conditions (again as shown by correlations between spectral tilt and fundamental frequency changes) while older adults with normal hearing thresholds and younger adults did not show this tendency.

In conclusion, investigating spontaneous speech in interaction in challenging communicative conditions can lead to a better understanding of the strategies used by speakers to maintain effective communication and of the impact of age and talker sex on such strategies. Despite the many challenges involved in the recording and analysis of spontaneous speech, such approaches will hopefully lead to a step forward in our knowledge of processes involved in speech communication.

References

Granlund, S., Hazan, V. L., & Mahon, M. (2015). Do children enhance phonetic contrasts in speech directed to a hearing-impaired peer? *Proceedings of the 18th International Congress of Phonetic Sciences*.

Hazan, V. L. & Baker, R. (2011). Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *Journal of the Acoustical Society of America* 130(4), 2139–2152

Hazan, V., Tuomainen, O. & Pettinato, M. (2016). Suprasegmental characteristics of spontaneous speech produced in good and challenging communicative conditions by talkers aged 9 to 14 years old. *Journal of Speech, Hearing and Language Research* 59, S1596–S1607.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In Hardcastle, W.J. & Marchal, A. (eds) *Speech Production and Speech Modelling*, The Netherlands: Kluwer Academic, 403–439.

Tuomainen, O. & Hazan, V. (2016). Articulation rate in adverse listening conditions in younger and older adults. *Proceedings of Interspeech 2016*, 8–12 September 2016, San Francisco, USA.

Van Engen, K. J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., & Bradlow, A. R. (2010). The Wildcat Corpus of Native and Foreign-Accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech* 53, 510–540.

"When two giants meet": The interaction between lexical stress and utterance-final prosody in spoken Hebrew

Vered Silber-Varod[1] and Noam Amir[2]
[1]The Research Center for Innovation in Learning Technologies, The Open University of Israel
[2]Department of Communication Disorders, Faculty of Medicine, Tel Aviv University

The present research investigates the interaction between the vowel duration at word level prosody and the phenomenon of phrase-final lengthening, as realized in spoken Israeli Hebrew. Experiments on the relations between the intonational phrase and lexical stress are numerous, and the phonological interaction of lexical and intonational features appears to be of typological interest (Ladd, 2001). Unlike experiments on how words in isolation are uttered, such studies strive to learn how phonological patterns are realized within natural environment, i.e., in an utterance or a sentence. For example, it is a common knowledge that syllables at breath group final position are longer than at its beginning (also known as *lengthening* and *anacrusis*, respectively (Cruttenden,1997)). Umeda (1975) studied such relations in large scale corpus of read speech. She found that the longer vowels were located at *pre-pausal* syllables, meaning, at stressed syllables by the end of utterance, paragraph, or sentence, to the rate of 1:2 or 1:3. The durational effects of phrase-final lengthening on the domain over which the speakers adjust it were studied in many languages, including Hebrew (Berkovits 1994). Turk and Shattuck-Hufnagel (2007) suggest two domains of the phrase-final lengthening in American English: the main stress syllable and the final syllable of the phrase-final word. They found that although most of the duration increase occurs in the phrase-final syllable rime, statistically significant lengthening of 7–18% also occurs in the main-stress syllable rime, when the main stress syllable *is not* the final syllable. Berkovits (1994) studied Hebrew stress manifestations of seven speakers, and found that utterance-final lengthening principally affected the final syllable regardless of stress. She also found evidence of progressive lengthening along the phrase-final syllable, which supports the suggestion that this phenomenon reflects the motor activity at utterance-final position, as the speech organs are slowing down. Silber-Varod and Kessous' (2008) study on weather broadcasts corpus, showed that stressed syllables in ultimate stress pattern located at intonation unit boundaries are longer than the preceding unstressed syllables, but stressed syllables in penultimate stress pattern are not always longer than the following unstressed utterance final syllables. The current database consists of 68 disyllabic target words in Hebrew, which differ phonemically only in their lexical stress pattern – final or penultimate. Target words were naturally embedded at the end of 68 carrier sentences (compared to six pairs of sentences and 12 disyllabic proper names in Berkovitz (1994)). Thirty subjects (13 Men and 17 Women) received the sentences in written form, and were instructed to read them aloud, pausing for at least two seconds between sentences. The research questions in the current study are: 1. How reliable is the contrast of the durational parameter between stressed and unstressed vowels of the same lexical word (not necessarily with identical vowel)?; 2. How reliable is the contrast of the durational parameter between stressed and unstressed identical vowels in the minimal pair words?

Results show that duration is an intrinsic indicator of stress, meaning the comparisons between stressed and unstressed vowels (p1* vs. p2, and u1 vs. u2* in Figure 1) of the same word showed significance differences ($t(29)=9.446$, -19.522 respectively, $p<0.001$ for both comparisons). As to the contrast between stressed and unstressed identical vowels (p1* vs. u1, and p2 vs. u2* in Figure 1), we found that duration is an extrinsic indicator of stress, i.e. the comparisons between stressed and unstressed identical vowels also showed significance differences ($t(29)=26.253$, 15.718 respectively, $p<0.001$ for both comparisons). Utterance-final lengthening affected the duration of the words and vowels. Words were ~26% longer at utterance-final position compared to nonfinal position. As to vowels, across words comparison showed that the effect on the final vowel was the largest, regardless of stress: Second vowels were lengthened to a larger extent (23% in penultimate stress and 18% in final stress) compared to first vowels (7% in penultimate stress and 4% in final stress). Within words comparisons showed a different effect on the two stress patterns: in penultimate words, the gap between stressed and unstressed vowels was reduced in utterance final position (35% in nonfinal; 17% in final position); in final stress words, the gap between stressed and unstressed vowels was *increased* in utterance final position (from 42% in nonfinal to 61% in final position). These findings suggest that although utterance final lengthening does not affect the relative dominancy in length of the stressed vowels, it lengthens the last vowels more than the first vowels, regardless of stress assignment.

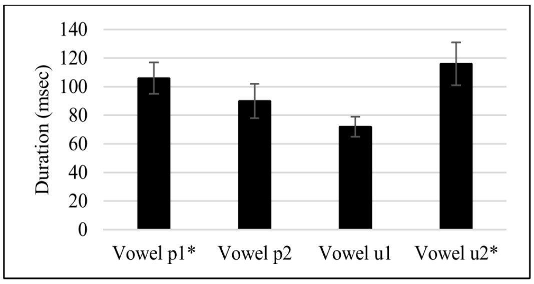

Figure 1. Mean vowel durations in utterance-final position, and 95% confidence intervals of four types of lexical stress conditions (p1*, p2, u1, and u2*). Asterisk [*] symbolizes the stress vowel.

Table 1. Means of durations (msec) for penultimate-stressed and final-stressed words in utterance-final position (current study, column B) compared to non-final position (Silber-Varod, Sagi, & Amir 2016, column A) and the study of Berokovits (1994) (columns D-F)

	A	B	C		D	E	F
	Non-final position (msec) (Silber-Varod, Sagi, & Amir 2016)	Utterance-final position (msec)	Difference		Non-final position (msec) (Berkovits, 1994)	Utterancefinal position (msec) (Berkovits, 1994)	Difference (Berkovits, 1994)
Penultimate stress							
Target word	337	424	26%		312	459	47%
Vowel p1*	99	106	7%		-	-	-
Vowel p2	73	90	23%		49**	77**	57%
Final stress							
Target word	336	427	27%		300	422	41%
Vowel u1	69	72	4%		-	-	-
Vowel u2*	98	116	18%		92**	127**	38%

* Stressed, ** In Berkovits (1994), target words had only [i] as the final vowel.

References

Berkovits, R. 1994. Durational effects in final lengthening, gapping and contrastive stress, *Language and Speech*, 37, 237–250.

Cruttenden, A. (1997). *Intonation.* Second edition. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Ladd, D. R. (2001). Intonational universals and intonational typology. In: M. Haspelmath (Ed.), *Language Typology and Language Universals: An International Handbook.* Vol. 2. De Gruyter, Berlin.

Silber-Varod, V., & Kessous, L. 2008. Prosodic boundary patterns in Hebrew: A case study of continuous intonation units in weather forecast. In P. A. Barbosa, S. Madureira,and C. Reis, (Eds.), *Proceedings of the Speech Prosody 2008 Conference*, Campinas, Brazil: Editora RG/CNPq. 265–268.

Silber-Varod, V., Sagi, H., & Amir, N. 2016. The acoustic correlates of lexical stress in Israeli Hebrew. *Journal of Phonetics*, 56, 1–14.

Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. Journal of Phonetics, 35(4), 445–472.

Umeda, N. (1975). Vowel duration in American English. *The Journal of the Acoustical Society of America*, 58(2), 434–445.

Electropalatographic analysis of vowels in quasi-spontaneous speech

Marko Liker
University of Zagreb, Croatia

Instrumental data on speech production physiology are crucial for our understanding of speech motor control, coarticulation and connected speech processes (Stone, 2010). However, instrumental physiological techniques, such as electropalatography (EPG), are mostly limited to laboratories and laboratory speech is often described as different from spontaneous speech (Celata and Calamai, 2012). It is therefore quite challenging to use instrumental physiological techniques for spontaneous speech recording and analysis. Particularly demanding for instrumental physiological analysis in both laboratory and spontaneous speech are vowels (Howard & Haselwood, 2012). Articulatory analysis of vowels is considered demanding mainly because of difficulties with determining tongue shapes and positions during speech. No single instrumental physiological technique provides a complete insight into vowel articulation and EPG might seem particularly unsuitable (Howard & Haselwood, 2012). EPG is instrumental physiological technique for recording and analysis of tongue-to-palate contact patterns during speech (Gibbon & Nicolaidis, 1999). Since tongue-to-palate contact is relatively low or even non-existent during vowel productions (Hardcastle & Gibbon, 1997), EPG is not the first choice when it comes to analysing vowels. However, some research data suggest that high vowels and some diphthongs can be analysed via EPG quite successfully (Byrd, 1995; Recasens & Espinosa, 2005; Gibbon et al. 2010). The analysis of low vowels using EPG is still problematic as well as separate quantification of vertical (high-low) and horizontal (front-back) position of each vowel using EPG data only.

Motivated by the review above, there are two aims in this investigation. The first aim is to elicit and record quasi-spontaneous speech in the laboratory using EPG. The second aim is to explore the possibilities of using EPG to quantify three Croatian corner vowels on the vertical (high-low) and horizontal (front back) axis of the vowel chart, which is comparable to the vowel chart based on acoustic data (F1, F2), in quasi-spontaneous speech.

Data were extracted from the R-kor corpus of Croatian speech containing simultaneous acoustic and EPG data. Speech material from eight female speakers of Standard Croatian with no speech or hearing impairments was utilised. A dialogue situation was set up in the form of a map task. Each speaker was asked to describe the path through a maze and read signs at 15 check-points marked throughout the path. Each sign contained a two-syllable CVCV word with one of the corner vowels of the Standard Croatian (/i, a, u/) in stressed and in unstressed position (e.g. "rasa", /'rasa/, English translation: race). During the recording session each speaker repeated each vowel five times in the stressed and unstressed position. Stressed vowels were analysed in this investigation. Six indices were calculated for each vowel: total contact, dental closure, alveolar closure, postalveolar closure, palatal closure and velar closure. Total contact was used to quantify the vertical (high-low) position of each vowel on the vowel chart. A newly developed measure used other five indices to quantify the horizontal (front-back) position of vowels. Articulate Assistant software (Wrench et al. 2002) was used for EPG analysis, while the statistical significance of differences was tested using two-way ANOVA with replication (alpha 0.05).

The clustering of the results for each vowel showed that on the basis of EPG indices it was possible to produce a vowel chart for each of the speakers. Two-way ANOVA with replication showed that differences between vowels were statistically significant both for vertical ($F(7, 2) = 660.24$, $p<0.001$) and horizontal ($F(7, 2) = 122.78$, $p<0.001$) axis of the vowel chart. The interaction analysis returned significant results (vertical: $p<0.001$, horizontal: $p=0.01$) showing that individual speakers' productions differed substantially. A newly developed measure for the quantification of vowels along the horizontal (front-back) axis proved to be more discriminative than the traditional CoG-based measure. The analysis of the variability of each of the indices showed that vowel /a/ is most variable in the majority of the speakers, while /i/ is least variable.

This investigation demonstrates one possibility of eliciting quasi-spontaneous speech in the laboratory using EPG. The investigation shows that EPG is sensitive enough to quantify vowels along vertical and horizontal axes of the vowel chart. The results of vowel variability are discussed in terms of Degree of Articulatory Constraint theory (Recasens et al. 1997).

References

Byrd, D. (1995). Palatogram reading a phonetic skill: A short tutorial. *Journal of the International Phonetic Association*, 24, 21–34.

Celata, C., Calamai, S. (2012). Introduction. *Italian Journal of Linguistics*, 24(1), 43–64.

Gibbon, F. E., Lee, A., Yuen, I. (2010). Tongue palate contact during selected vowels in normal speech. *The Cleft Palate Craniofacial Journal*, 47(4), 405–412.

Gibbon, F. E., Nicolaidis, K. (1999). Palatography. In W. J. Hardcastle, N. Hewlet (eds.) *Coarticulation: theory, data and techniques*, Cambridge: CUP, 229–245.

Hardcastle, W. J., Gibbon, F. (1997). Electropalatography and its clinical applications. U: M. J. Ball, & C. Code (ur) *Instrumental Clinical Phonetics*. Whurr: London.

Howard, S., Haselwood, B. (2012). The contribution of phonetics to the study of vowels and vowel disorders. U: Ball, M. J., Damico, J. S., Gibbon, F. E. (ur) *Handbook of Vowels and Vowel Disorders*. Routledge: London, 61–112.

Recasens, D., Espinosa, A. (2005). Dispersion and variability of Catalan vowels. *Speech Communication*, 48(6), 645–666.

Recasens, D., Pallarès, M. D., Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America* 102 (1): 544–561.

Stone, M. (2010). Laboratory techniques for investigating speech articulation. In W. J. Hardcastle, J. Laver, F. E. Gibbon (eds.). *The Handbook of Phonetic Sciences*. Malden-Oxford-Chichester: WileyBlackwell. 9–38.

Wrench A. A., Gibbon, F. E., McNeill, A. M., Wood, S. E. (2002). *An EPG therapy protocol for remediation and assessment of articulation disorders*. In John H.L. Hansen; Brian L. Pellom (Eds.). *Proceedings of ICSLP-2002*: 965–968.

Disfluencies in spontaneous speech in younger and older adults in easy and difficult communicative situations

Outi Tuomainen and Valerie Hazan
Speech Hearing and Phonetic Sciences, University College London (UCL), UK

Spontaneous conversational speech is notoriously disfluent (Bortfeld et al., 2001): disfluencies (DFs) such as word and phrase repetitions and false starts may form even up to 5-10% of everyday conversations (Clark, 1994). Disfluencies in spontaneous speech are often associated with disruptions in word-finding or formulating sentences, with distractions (Yairi & Seery, 2011) or with an increase in cognitive load (Bortfield et al., 2001). It has also been shown that talkers become more disfluent when they are speaking in background noise (Jou & Harris, 1992, Southwood & Dagenais, 2001). There are well-documented changes in speech perception and production with increasing age. Older talkers have more difficulty retrieving words than do younger talkers (Burke et al., 1991) generating more disfluencies in discourse compared to younger adults (YAs) (Bortfeld et al., 2001). Older adult (OA) talkers also often report having particular difficulty communicating in challenging listening conditions, e.g. in noise or in the presence of other talkers.

When communication becomes effortful, talkers need to make various adjustments to their speech production to aid listener's understanding. For example, they may use a 'clear speaking style', which involves them speaking more slowly and reducing the complexity of their utterances amongst other characteristics. Adopting a more careful speaking style is likely to result in decrease in disfluencies in conversation, but it is not known whether OA talkers become less disfluent when speaking clearly than casually. Moreover, it is not known whether OA talkers become more disfluent than YA talkers when they are communicating in challenging listening conditions. The aim of the current study was to investigate disfluency rates in younger and older adults when they are speaking casually, when they are speaking clearly for the benefit of their interlocutor, and when they are speaking in background noise.

83 older adults aged 65 to 85 years (30 female) and 26 younger adults aged 18 to 35 years (15 female) were recorded while they completed a problem-solving spot-the-difference picture task (diapixUK; Baker & Hazan, 2011) with a young adult interlocutor (aged 18-33 years). The main participants (OA, YA) were told to take the lead in the interaction ('Talker A' role) while the young adult interlocutor had a more passive role ('Talker B'). Talker pairs completed the tasks in three different listening conditions: when no interference was present (NORM), when Talker B had a simulated severe-to-profound hearing loss (HLS), and when both talkers heard 8-talker babble noise (BAB2). It was expected that the NORM condition would elicit a casual speaking style in Talker A while the HLS and BAB2 would elicit a clear speaking style, as this would be necessary to communicate effectively despite the communication barrier. DFs were classified from Talker As speech using a system adapted from Shriberg's Disfluency Types (Shriberg, 2001) that has previously been used to analyse spontaneous speech (see Table 1). DF types and their position in an utterance (not reported here) were identified manually in Praat (version 6.0.19) and their frequencies were extracted using an in-house Praat script. Because the length of the speech samples differed between different speakers, the disfluency rate was calculated as a percentage of disfluent items relative to the total number of words produced in each listening condition. We predicted that all talkers would produce more disfluencies in a casual speaking style than in a clear style. We also predicted that OA talkers would produce more disfluencies overall than YA talkers both when communication was easy (NORM) and difficult (HLS and BAB2), and that they would be particularly affected by the background babble. In addition to the effects of age, some studies have shown that men produce more disfluencies than women (Shriberg, 1994; Bortfield, 2001), and we predicted that we would find these gender differences across all speaking styles.

Preliminary statistical results based on an analysis of a subset of talkers (N=20) across all disfluency types show that, as predicted, OA talkers produced more disfluencies (7.8%) than YA talkers (6.2%), and male talkers (8.3%) more than female talkers (5.5%) in the NORM condition. Furthermore, female talkers (both YA and OA) produced less disfluencies when they were talking clearly for the benefit of their interlocutor (HLS condition, see Table 2). However, male talkers did not show the same disfluency reduction in difficult communicative conditions. Against our predictions, when talking in background noise (BAB2), older adult male talkers produced less disfluencies than when communicating in good listening conditions (NORM). The other talker groups did not significantly increase or decrease the disfluencies in background noise (see Table 2). However, descriptive data (and preliminary statistical analyses) show that older adult female talkers produced marginally more disfluencies in the BAB2 condition than in the NORM condition indicating a potential difficulty communicating in background noise. Together these preliminary results indicate that there are age- and gender-related differences in disfluency rates in casual spontaneous speech. Furthermore, there are potential age and gender differences when talkers are either adapting their speech for the benefit of their interlocutor or communicating in noise. However, these results are based on a subset of the data and should be treated as showing possible statistical trends. Analysis from a larger set of talkers will be presented at the workshop. Analyses of other factors (such as speaking rate and hearing status) that might affect disfluency rates in these groups, along with analyses of different disfluency types (including within-speaker silent pauses) will also be presented.

Table 1. Disfluency types adapted from Shriberg, 2001.

Group of disfluency	Type of disfluency	Example
Filled pauses	Erm, err Other	Show flights from Boston on (erm) from Denver on Monday Show flights from (like) Boston
Repetitions	Word repetitions Phrase repetitions Part-word repetitions Insertions Articulation errors Substitutions	Show the – the morning flights Show the – show the morning flights Show the morn – morning flights Show the flights – show the morning flights Show me the flee – flights Show the morning – show the evening flights
Incomplete phrases	False starts	Show me the – which flights go to Boston