I just thought like: The phonetic variation of *just* with applications to forensic voice comparison

Ben Gibb-Reid, Paul Foulkes, Vincent Hughes, and Traci Walker
University of York / University of Sheffield

A specific difficulty of the task of forensic voice comparison (FVC) is the lack of direct correspondence in the content of different recordings. That is, recordings are unlikely to share many of the same words. Therefore, a frequently used word (or other feature) in naturally occurring speech may be of value to the FVC practitioner. To examine the usefulness of a feature in FVC, it is necessary to understand how variable it is between speakers, and the factors that affect it in different discourse positions or prosodic contexts. A short discourse-pragmatic marker (DPM) such as *just* could be analysed in this way for suitability as a diagnostic feature in FVC. In previous research, other DPMs such as filled pauses (*uh, um*) have been analysed as forensic voice comparison features with promising results (Foulkes, Carrol & Hughes, 2004; Tschäpe et al., 2005; Hughes, Wood & Foulkes, 2016).

Studies of the DPM *just* have focused on its various discourse functions and overall frequency of occurrence. Aijmer (2002; 2005), Beeching (2016) and Woolford (2021) highlight adverbial functions of *just* ('to be precise' and 'to reference the recent past') as distinct from emphatic (restricting / intensifying meaning) or planning (when used as a 'filler') functions. Phonetic variation of *just* has not been the subject of detailed study (with the exception of the Canberra Corpus Collective, forthcoming). This study compares variation in both the discourse functions of *just* and the phonetic realisation of its segments, in a similar way to analyses applied to *like* by Drager (2011) and Schleef and Turton (2016).

In the spoken data of *The British National Corpus* (2007; BNC), *just* is the 42nd most frequent word, occurring 0.39 times per 100 words (or 3,850 times per million words). In a preliminary analysis of data from 10 male Southern Standard British English speakers (DyViS corpus; Nolan, McDougall, De Jong & Hudson, 2009), *just* is, as predicted, highly frequent in informal talk. In a count of all occurrences of the word, including both adverbial and discourse-pragmatic tokens, *just* had a frequency of 0.8 per 100 words, i.e., approximately double the BNC estimate. There is also an indication that the frequency of *just* varies markedly between speakers – a range from 0.33 to 1.06 per 100 words. Such variation is exactly what is needed for a good diagnostic feature in FVC. More data will be collected to confirm this observation, and to explore whether *just* varies in function or prosodic context across speakers.

Analyses of *just* will describe the realisation of its four segments (in citation form /d͡ʒ ʌ s t/) including vowel formant data, in respect of both functional categories and prosodic contexts. Data will be collected from all 100 DyViS speakers, which will enable us to run some speaker discriminant tests to assess accuracy for FVC suitability. There will also be a comparison of the effectiveness of the vowel in *just* relative to filled pauses and STRUT vowels from the same speakers in a similar analysis previously undertaken in FVC research (Foulkes et al., 2004).

References

Aijmer, Karin. 2002. English discourse particles: Evidence from a corpus. Amsterdam: John Benjamins.

Aijmer, K. (2005). 'Just' and multifunctionality. Contexts - Historical, Social, Linguistic: Studies in Celebration of Toril Swan. K. McCafferty, Bull, Tove & Killie, Kristin, Peter Lang: 31-47.

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

The British National Corpus, version 3 (2007). Distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/

CCC (forthcoming) It's not just a sound change: linking phonetic and pragmatic change in a discourse-pragmatic marker.

Drager, K. K. (2011). "Sociophonetic variation and the lemma." Journal of Phonetics 39(4): 694-707.

Foulkes, P., Carrol, G. & Hughes, S. (2004) Sociolinguistics and acoustic variability in filled pauses. Paper presented at the annual conference of the International Association for Forensic Phonetics and Acoustics, Helsinki, Finland.

Gold, E., Ross, S. & Earnshaw, K. (2018). The 'West Yorkshire Regional English Database': Investigations into the generalizability of reference populations for forensic speaker comparison casework. Interspeech 2018: Speech Research for Emerging Markets in Multilingual Societies.

Hughes, V., Wood, S. & Foulkes, P. (2016). Strength of forensic voice comparison evidence from the acoustics of filled pauses. International Journal of Speech, Language & the Law. 23.

Nolan, F., McDougall, K., De Jong, G. & Hudson, T. (2009). "The DyViS database: style-controlled recordings of 100 homogeneous speakers for forensic phonetic research." International Journal of Speech Language and the Law 16(1): 31-57.

Schleef, E. & Turton, D. (2016). "Sociophonetic variation of like in British dialects: effects of function, context and predictability." English Language and Linguistics 22(1): 35-75.

Tschäpe, N., Trouvain, J., Bauer, D., & Jessen, M. (2005). Idiosyncratic patterns of filled pauses. Paper presented at the 14th Annual Conference of the International Association for Forensic Phonetics and Acoustics (IAFPA), Marrakesh, Morocco.

Woolford, K. (2021). Just in Tyneside English. World Englishes. doi:10.1111/weng.12542