Disfluency patterns in the language production system

Aurélie Pistono, Robert J. Hartsuiker

Department of experimental psychology, Ghent University, Belgium

Abstract

Within the language system, several of the language production levels may be involved in the production of disfluencies. Here, we conducted network task experiments to tackle disfluencies occurring during lexical selection, grammatical selection, and conceptual formulation. We showed that each difficulty induced a different pattern of disfluency. Additionally, multivariate pattern analyses demonstrated that difficulty is predictable from disfluency data patterns.

Keywords: language production, disfluency, multivariate pattern analyses

Introduction

Natural speech production is full of disfluencies. This term includes various phenomena such as filled or silent pauses, repeated words, and self-corrections. However, it is still not well understood why speakers are so often disfluent. To reveal the underlying cause of these phenomena, several authors attempted to relate the pattern of disfluencies to difficulties at specific levels of production, using network tasks (e.g., Hartsuiker and Notebaert, 2010; Figure 1). In this paradigm, participants describe a route through a network of pictures. This allows for the manipulation of the items to create difficulties at specific stages (e.g. conceptual generation) while holding others constant (e.g. lexical selection). Here, we conducted two experiments to examine the pattern of disfluency related to lexical selection difficulty, grammatical selection difficulty, and conceptual difficulty. Additionally, we aimed at examining whether, by contrast, the manipulated difficulty could be predicted based on the pattern of disfluency associated with it, using multivariate pattern analyses (MVPA, Haynes & Rees, 2006). Instead of analysing each dependent variable individually, MVPA extracts the information contained in the pattern of information available, to test whether experimental conditions can be distinguished from one another on the basis of the patterns observed.

Material and methods[1]

In each experiment, 20 bachelor students, all native Dutch speakers performed 20 network tasks. Pictures were either connected by one, two, or three straight lines or curves, that had a fixed length. The route through the network was indicated by a moving red dot that traversed the network in 42 seconds.

Instructions were given to provide an accurate description of the network using complete sentences and to synchronize the description with the dot that moved through the network.

In *Experiment 1*, we manipulated pictures' name agreement to examine the initial stage of lexical access. We also examined grammatical selection through grammatical gender. 160 pictures were selected from Severens et al., 2005 (i.e., eighty pictures had high name agreement and eighty had low name agreement; eighty pictures had a common gender name and eighty had a neuter gender name). Pictures were matched for their log frequency and age of acquisition. In *Experiment 2*, we examined the conceptual generation of the message, by manipulating the visual identification of some items (i.e., by blurring). We selected 160 pictures from the Multipic database (Duñabeitia et al., 2018), that we controlled for name agreement, age of acquisition, and visual complexity. For each participant, eighty pictures had a four pixel radial blur and eighty were control pictures.
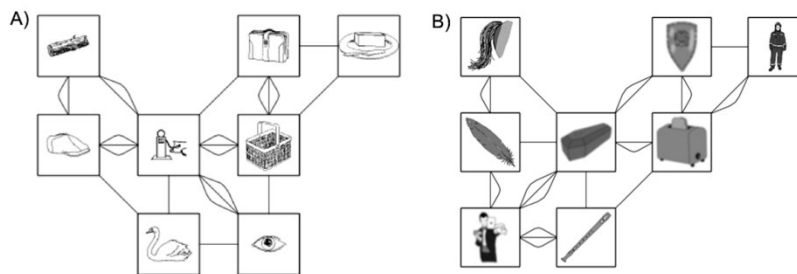

Figure 1. Example of a network for A) Experiment 1 and B) Experiment 2

All productions were transcribed and scored by a native Dutch speaker and checked by another native speaker. We analysed self-corrections, silent pauses, filled pauses, and prolongations related to each picture. In each experiment, we ran linear-mixed effects models with name agreement (low/high), grammatical gender (neuter/common), or conceptual difficulty (blurred/non-blurred items) as a fixed effect. Additionally, we used MVPA to predict whether each participant was about to mention a low or high name agreement item, a common or neuter gender item, or a blurred or non-blurred item. We trained a linear discriminant analysis classifier on the four disfluency features, using the Scikit-learn toolbox (Pedregosa et al., 2011).

Results

In Experiment 1, 16.3% of pictures induced at least one disfluency. Low name agreement items induced more self-corrections ($\chi^2(1)=10.19$, $p<.01$), silent pauses ($\chi^2(1)=17.59$, $p<.0001$), and filled pauses ($\chi^2(1)=11.32$, $p<.001$) than high name agreement. Common gender items elicited more prolongations

($\chi^2(1)$=16.21, p<.0001) than neuter gender items. The classifier could predict from the pattern of disfluency whether a participant was about to name low or high name agreement items (59.57% correct on average; t(19)=7.04, p<.001). The contribution of each feature was consistent across participants for this classification (self-corrections: (t(19)=3.6, p<.01); silent pauses: (t(19)=6.5, p<.0001); prolongations: (t(19)=3.2, p<.01); filled pauses: t(19)=3.5, p<.01). The classifier could also predict items' gender (53.72% correct on average; t(19)=3.83, p=.001). Only prolongations' contribution was consistent (t(19)=-2.5, p<.05).

Experiment 2 elicited more disfluencies than Experiment 1 (26% of pictures). Surprisingly, blurry pictures did not elicit more disfluency than control pictures. Classification accuracies were above chance level (53.24%; t(18)=2.77, p<.01) but the contribution of disfluencies was not consistent, suggesting inter-individual variability in disfluency patterns (Figure 2).
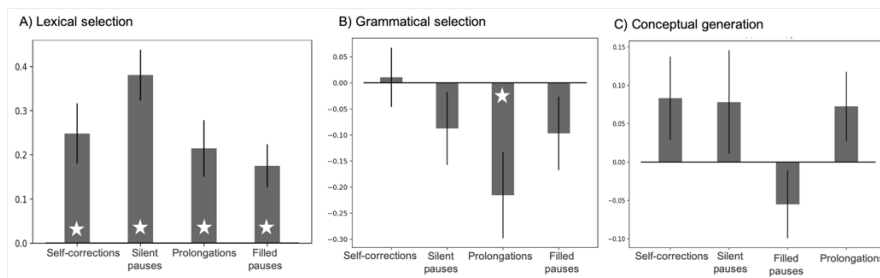

Figure 2. Contribution of each feature when classifying the pattern of disfluency related to each manipulation. White stars indicate significance.

Discussion

Difficulties in determiner selection led to a different pattern of disfluencies than difficulties in content word selection. This suggests that difficulties at distinct stages of production elicit different disfluencies. The finding that low name agreement induces pauses may indicate that they reflect an 'act of choice' between lexical items with similar semantic features (Beattie & Butterworth, 1979). Because of this 'act of choice', speakers are also more error-prone, leading to self-corrections. Surprisingly, common gender induced more disfluency than neuter gender. Possibly, this effect is related to the phonological form of this determiner ('de' in opposition to the neuter gender determiner 'het'), which is more likely to encourage prolongations.

In Experiment 2, impeding the visual identification of some items influenced the whole task, by increasing disfluency overall. It is possible that making the identification of some items more difficult might have hampered the monitoring system (i.e., poor error detection and correction while staying synchronized with the pace of the dot) throughout the whole task, similarly to time pressure (Oomen & Postma, 2001). MVPA provided complementary

findings. They revealed that the pattern of disfluencies is informative about the linguistic difficulty of an item, i.e. a classifier was able to learn and predict the type of item a speaker is about to name. For the conceptual manipulation in particular, it clarifies why linear mixed models were not able to reveal significant differences: conceptual difficuly manifests itself differently from one participant to another. These analyses represent a further step for current models of language production, to capture inter-individual variability.

Notes

1. Details about the project and methods are available here: https://osf.io/9yhcb/

Acknowledgements

References

Beattie, G.W., Butterworth, B.L. 1979. Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. Language and Speech, 22(39), 201–211.

Duñabeitia, J. et al. 2018. MultiPic: A standardized set of 750 drawings with norms for six European languages. Quarterly Journal of Experimental Psychology, 71(4), 808–816.

Hartsuiker, R.J., Notebaert, L. 2010. Lexical access problems lead to disfluencies in speech. In Experimental Psychology. 57(3), 169–177.

Haynes, J.-D., Rees, G. 2006. Decoding mental states from brain activity in humans. Nature Reviews Neuroscience, 7(7), 523–534.

Oomen, C.C., Postma, A. 2001. Effects of time pressure on mechanisms of speech production and self- monitoring. J Psycholinguist Res, 30(2), 163–184.

Pedregosa, F., et al. 2011. Scikit-Learn: Machine Learning in Python. J. Mach. Learn. Res., 12, 2825–2830.

Severens, E., et al. 2005. Timed picture naming norms for 590 pictures in Dutch. Acta Psychologica, 119(2), 159–187.