

The role of co-occurrence patterns in the acquisition of sociolinguistic variation: shaping a methodological framework

Moira Van Puyvelde¹, Laura Rosseel¹, Eline Zenner² & Dirk Speelman²
¹Vrije Universiteit Brussel, Moira.Van.Puyvelde@vub.be ²KU Leuven

Keywords: Cognitive Sociolinguistics, Language acquisition, Sociolinguistic variation, Artificial language learning

BACKGROUND | The usage-based thesis assumes that language acquisition results from the extraction of linguistic units from input (Langacker 1987, Tomasello 2003, Bybee & Beckner 2012). Frequency is central to that process: the more often an utterance is encountered, the more entrenched it becomes in memory (Langacker 2000, Boyland 2009). From a cognitive sociolinguistic perspective, it has been argued that the frequency with which language variants and social categories co-occur in input influences the speed with which the mental link between them is acquired by language users (Foulkes 2010; Docherty, Langstrof & Foulkes 2013). However, experimental research on this topic is limited in number and methodologically diverse: the differing design choices from previous studies complicate the construction of a uniform framework to answer the fundamental question of how socially meaningful language variation is learned.

RESEARCH AIMS | The aim of the present study is twofold. First, it examines how the distributional characteristics of linguistic input relate to the acquisition of sociolinguistic variation. Second, the study proposes a coherent methodological paradigm to experimentally investigate this and related developmental questions of socially meaningful language variation by building on and optimising design choices from earlier work in a structured way.

METHOD | Our experiment employs semi-artificial language learning to train participants on a non-existing sociolinguistic variable, thereby exploring various co-occurrence patterns of language variants and social categories in a fixed amount of input. A sample of 80 Flemish students is exposed to short sentences in Dutch that contain a pseudonoun (e.g. Ik zie een stipis. "I see a stipis."). This pseudoword can contain either one of two different variants, viz. a voiced versus voiceless medial stop, which co-occur probabilistically with two speaker gender identity groups, i.e. speakers identifying as male versus as female. Co-occurrence patterns of variants and speaker groups are manipulated between participants: one variant is used by one group in 90%, 75% or 60% of the group's total number of utterances. After exposure, participants complete a categorisation task that measures their learning of associations between the linguistic variants and speaker gender identity, also testing for pseudowords not encountered in the training phase.

EXPECTED RESULTS | Frequency of co-occurrence strengthens the mental representation of utterances and facilitates the retrieval of linguistic and social information (Boyland 1996, Bybee & Thompson 2000, Ibbotson 2013, Drager & Kirtley 2016). We therefore expect more frequent co-occurrence of the same linguistic variant with the same social setting to result in more successful learning across participants, i.e. in the 90% condition followed by the 75% and 60% conditions respectively. Our results will contribute both to the refinement of models of the cognitive representation of sociolinguistic knowledge and to the development of a methodological paradigm to address the question of how listeners acquire the social meaning of linguistic variation.

References

- Boyland, Joyce T. (1996). *Morphosyntactic change in progress: A psycholinguistic treatment*. Berkeley, CA: University of California dissertation.
- Boyland, Joyce T. (2009). Usage-based models of language. In David Eddington (ed.), *Experimental and quantitative linguistics* (pp. 351-419). Munich: Lincom.
- Bybee, Joan L. & C. Beckner. (2012). Usage-based theory. In Heine, Bernd & H. Narrog (eds.), *The Oxford handbook of linguistic analysis* (pp. 827-856). Oxford: Oxford University Press.
- Bybee, Joan L. & S. Thompson. (2000). Three frequency effects in syntax. *Berkeley Linguistics Society* 23(1). 378-388.
- Docherty, Gerard, C. Langstrof & P. Foulkes. (2013). Listener evaluation of sociophonetic variability: Probing constraints and capabilities. *Linguistics* 51(2). 355-380.
- Drager, Katie & J. Kirtley. (2016). Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. In Babel, Anna M. (ed.), *Awareness and control in sociolinguistic research* (pp. 1-24). Cambridge: Cambridge University Press.
- Foulkes, Paul. (2010). Exploring social-indexical knowledge: A long past but a short history. *Laboratory Phonology* 1(1). 5-39.
- Ibbotson, Paul. (2013). The scope of usage-based theory. *Frontiers in Psychology* 4. Article 255.
- Langacker, Ronald W. 1987. *Foundations of cognitive grammar. Vol. 1: Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. (2000). A dynamic usage-based model. In Barlow, Michael & S. Kemmer (eds.), *Usage-based models of language* (pp. 1-64). Stanford, CA: CSLI Publications.
- Tomasello, Michael. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge & London: Harvard University Press.