

The lexicon adapts to interaction: Pragmatic motivation for systematic sound symbolism in *wh*-words

Anita Slonimska and Sean G. Roberts

(CLS, Radboud University/Max Planck Institute for Psycholinguistics)

The aim of this study was to investigate whether there is a universal tendency for content interrogative words (*wh*-words) within a language to sound similar in order to facilitate pragmatic inference in conversation.

One of the challenges of everyday conversation is the timing of turn-taking. Gaps between turns are very short, only about 200ms, (Stivers et al., 2009), meaning that listeners must begin planning their turn as soon as possible. Previous literature has suggested several types of cue that aid the recognition of speech-acts such as prosody (Sicoli et al., 2014) and eye gaze (Rossano, Brown & Levinson, 2009). We hypothesise that a systematic phonetic cue that marks interrogative words would also help early recognition of questions (allowing early preparation of answers), for instance initial segments of *wh*-words sounding similar within a language.

We analyzed 173 languages from 65 different language families by means of permutation tests. The similarity of initial segments of *wh*-words was quantified by means of entropy efficiency score. We found that initial segments of *wh*-words were more similar within a language than expected by chance, also when controlling for language family, geographic area (stratified permutation) and analyzability (compound phrases excluded).

- (1) Phonemic transcriptions of *wh*-words of two languages that have similar initial segments (low entropy) and one language that has dissimilar initial segments (high entropy)

Wh-word	Bulgarian (Indo-European)	Aymara (Aymaran)	Dehong (Tai-Kadai)
how	<i>kak</i>	<i>kumasa</i>	<i>comsə</i>
how many	<i>kolko</i>	<i>kawkanaksa</i>	<i>xo</i>
how much	<i>kolko</i>	<i>kawksa</i>	<i>lalajomlai</i>
what	<i>kakvo;sto</i>	<i>kunasa</i>	<i>ʔen</i>
when	<i>koga</i>	<i>kunarsa</i>	<i>hakkek</i>
where	<i>kəde</i>	<i>kawkinsa</i>	<i>cup</i>
which	<i>koj</i>	<i>kawkisa</i>	<i>həpʔum</i>
who	<i>koj</i>	<i>kitisa</i>	<i>xincep</i>
why	<i>zasto</i>	<i>kunatsa</i>	<i>maicai</i>

Next, we investigated whether similarity of initial segments of question words was larger than initial segments of random and conceptually related words, reasoning that for a pragmatic function in interaction to be effective the category of question words should stand out on the background of the content information. Test of random samples revealed that initial segments of *wh*-words were more similar than initial segments of conceptually related sets of words (e.g., body parts) and random word sets.

Finally, we hypothesized that this cue would be more useful at the beginning of a turn, considering that no paralinguistic cues are yet available for the addressee to recognize a speech-act as a question; so similarity of the initial segment of *wh*-words should be greater in languages that place them at the beginning of a clause. We gathered typological data on 95

languages, and found the predicted trend, although statistical significance was marginal when controlling for historical contact. Importantly, we also found that the pattern was reversed for random words; namely, random words tended to be more diverse in languages that use initial interrogative phrase in comparison to languages that do not.

While there may be several mechanisms that bring about this pattern (e.g., common derivation), we suggest that the ultimate explanation of the similarity of *wh-words* is to facilitate early speech-act recognition. Importantly, this hypothesis can be tested empirically, and the current results provide a sound basis for future experimental tests.

References

- Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, questioning and culture. *Conversation analysis: Comparative perspectives*, 187-249.
- Sicoli, M. A., Stivers, T., Enfield, N. J., & Levinson, S. C. (2014). Marked initial pitch in questions signals marked communicative function. *Language and Speech*, 0023830914529247.
- 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. *Proceedings of the National Academy of Sciences*, 106(26), 10587- 10592.