Paradigmatic effects in Estonian inflected noun production

Kaidi Lõo[1]; R. Harald Baayen [2] [1] Fabian Tomaschek[2]; Benjamin V. Tucker [1]; Juhani Järvikivi[1];

[1]University of Alberta, Canada; [2] University of Tübingen, Germany

BACKGROUND

- It is generally assumed that majority of inflected forms are processed as discrete morphemic constituents (Levelt et al., 1999; Pinker, 1999).
- Recent studies show that also whole-word frequency (Baayen et al., 1997), inflectional entropy (Milin et al., 2009) and morphological family size (De Jong et al., 2002) co-determine lexical processing costs.

CURRENT STUDY

- The rich inflectional morphology of Estonian offers the possibility to investigate a further predictor, **inflectional paradigm size**, the number of actually attested forms of a given paradigm.

Table 1: Inflectional paradigm of *jalg* 'foot, leg' with 46 paradigm members. The 36 paradigm members present in the corpus are marked in bold. Case

Case	Singular	Plural	Translation
Nominative	**jalg**	**jalad**	foot (subject)
Genitive	**jala**	**jalgade**, **jalge**	of a foot
Partitive	**jalga**	**jalgasid**, **jalgu**	foot (object)
Illative-1	**jalga**	-	into a foot
Illative-2	jalasse	**jalgadesse**, jalusse, jalgesse	into a foot
Inessive	**jalas**	**jalgades**, **jalus**, **jalges**	in a foot
Elative	**jalast**	**jalgadest**, **jalust**, **jalgest**	from a foot
Allative	**jalale**	**jalgadele**, **jalule**, **jalgele**	onto a foot
Adessive	**jalal**	**jalgadel**, **jalul**, **jalgel**	on a foot
Ablative	**jalalt**	**jalgadelt**, jalult, **jalgelt**	from a foot
Translative	**jalaks**	**jalgadeks**, jaluks, jalgeks	turn into a foot
Terminative	jalani	**jalgadeni**, jalgeni	up to a foot
Essive	jalana	jalgadena	as a foot
Abessive	**jalata**	**jalgadeta**	without a foot
Comitative	**jalaga**	**jalgadega**	with a foot

RESEARCH QUESTION

- Do inflected form's own **frequency** and actual **inflectional paradigm size** affect the production of Estonian case-inflected nouns?

METHODOLOGY

- **Experiment 1**: 26 native speakers of Estonian (18 female; 21-67 years) reading aloud 200 Estonian case-inflected nouns.
- **Experiment 2**: 33 native speakers of Estonian (20 female; 22-69 years) reading aloud 2,800 Estonian case-inflected nouns.

ANALYSIS

- Generalized additive and quantile regression (Wood, 2006; Fasiolo et al., 2016).
- **Response variable**: (1) production latency; (2) acoustic duration
- **Predictors**: (1) whole-word frequency, e.g. **works**
  (2) inflectional paradigm size, e.g. **work**, **works**; **worked**
  (3) morphological family size, e.g. **worker**, **workforce**, **handwork**
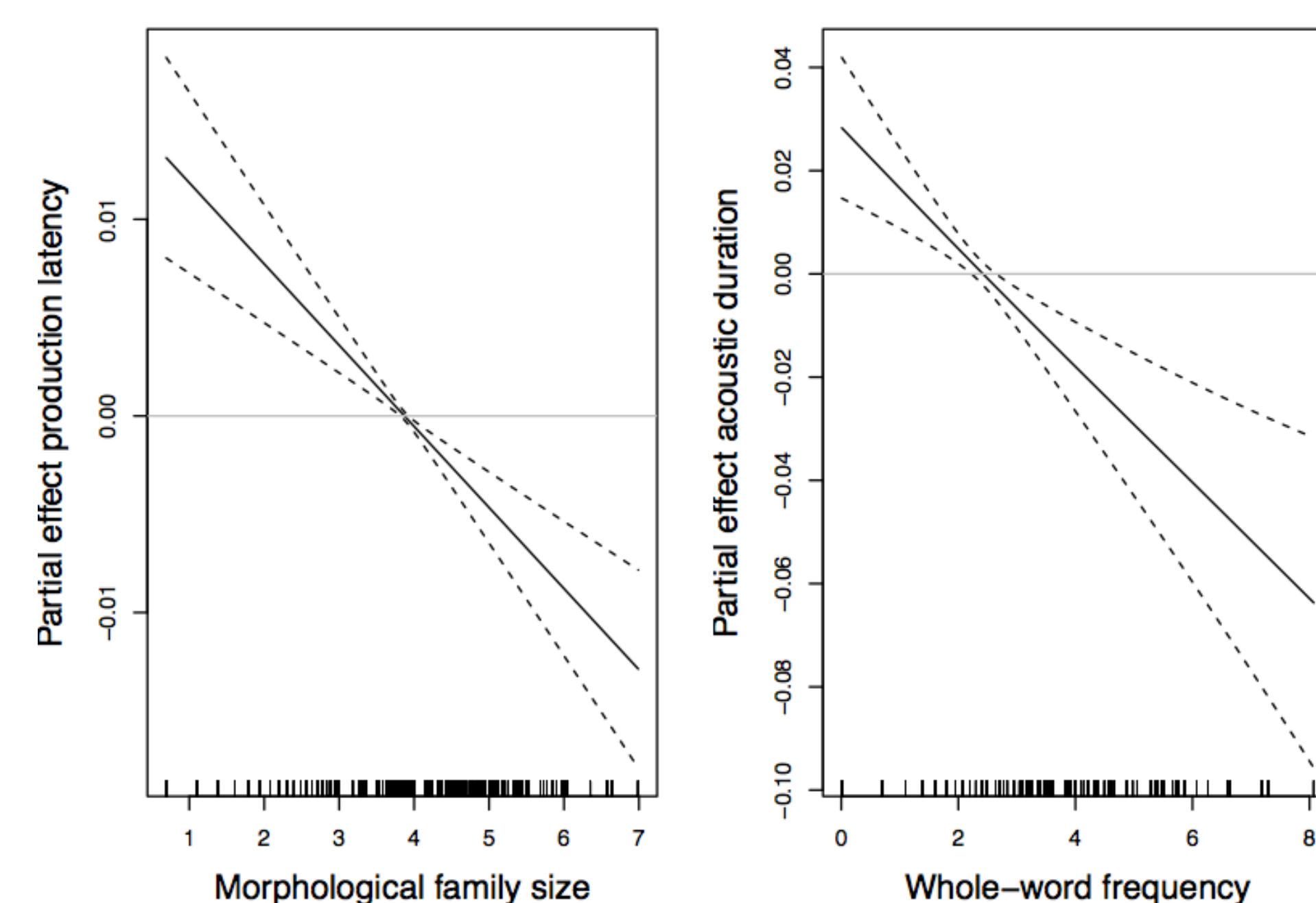
RESULTS

Figure 1: Partial effects for morphological family size and whole-word frequency in Experiment 1.
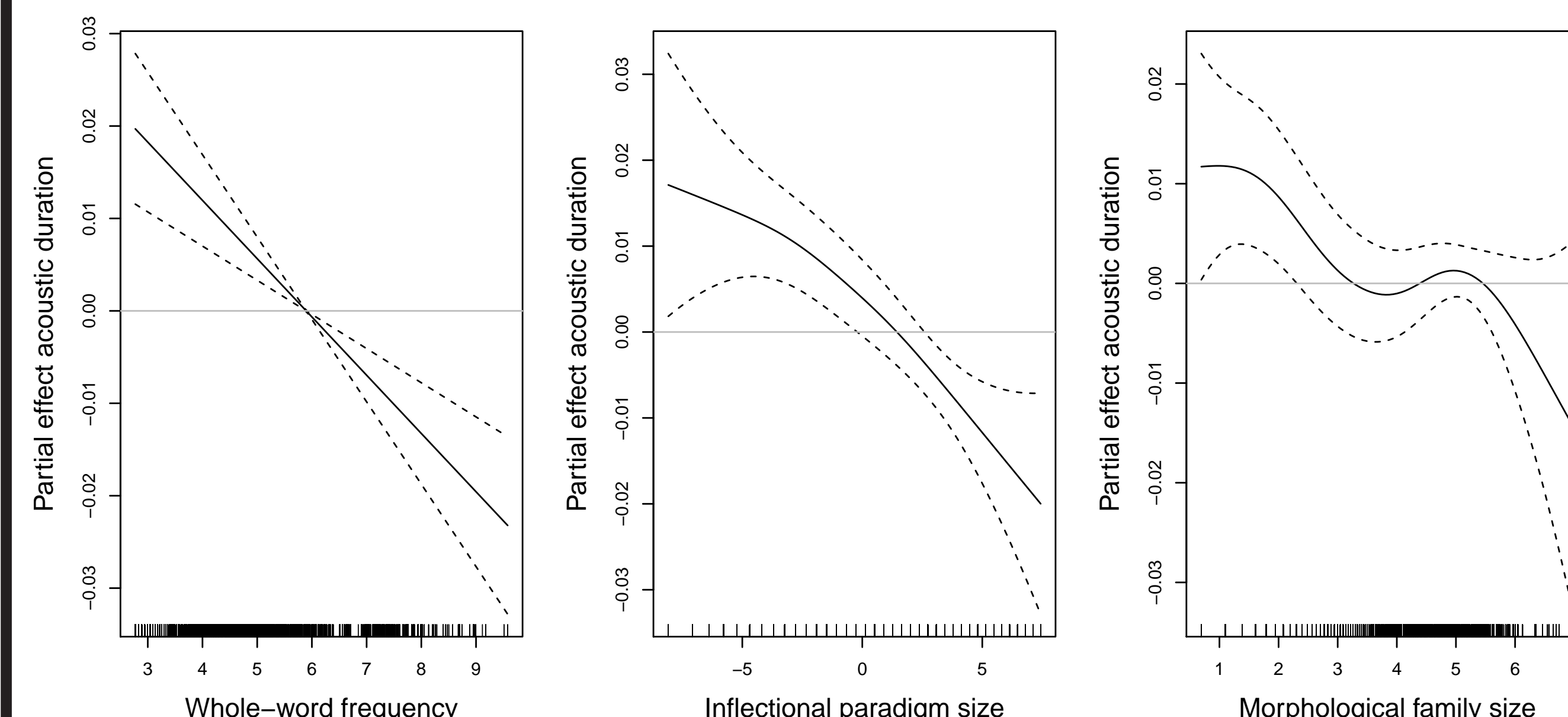

Figure 2: Partial effects for whole-word frequency, inflectional paradigm size, morphological family size in acoustic duration of Experiment 2.
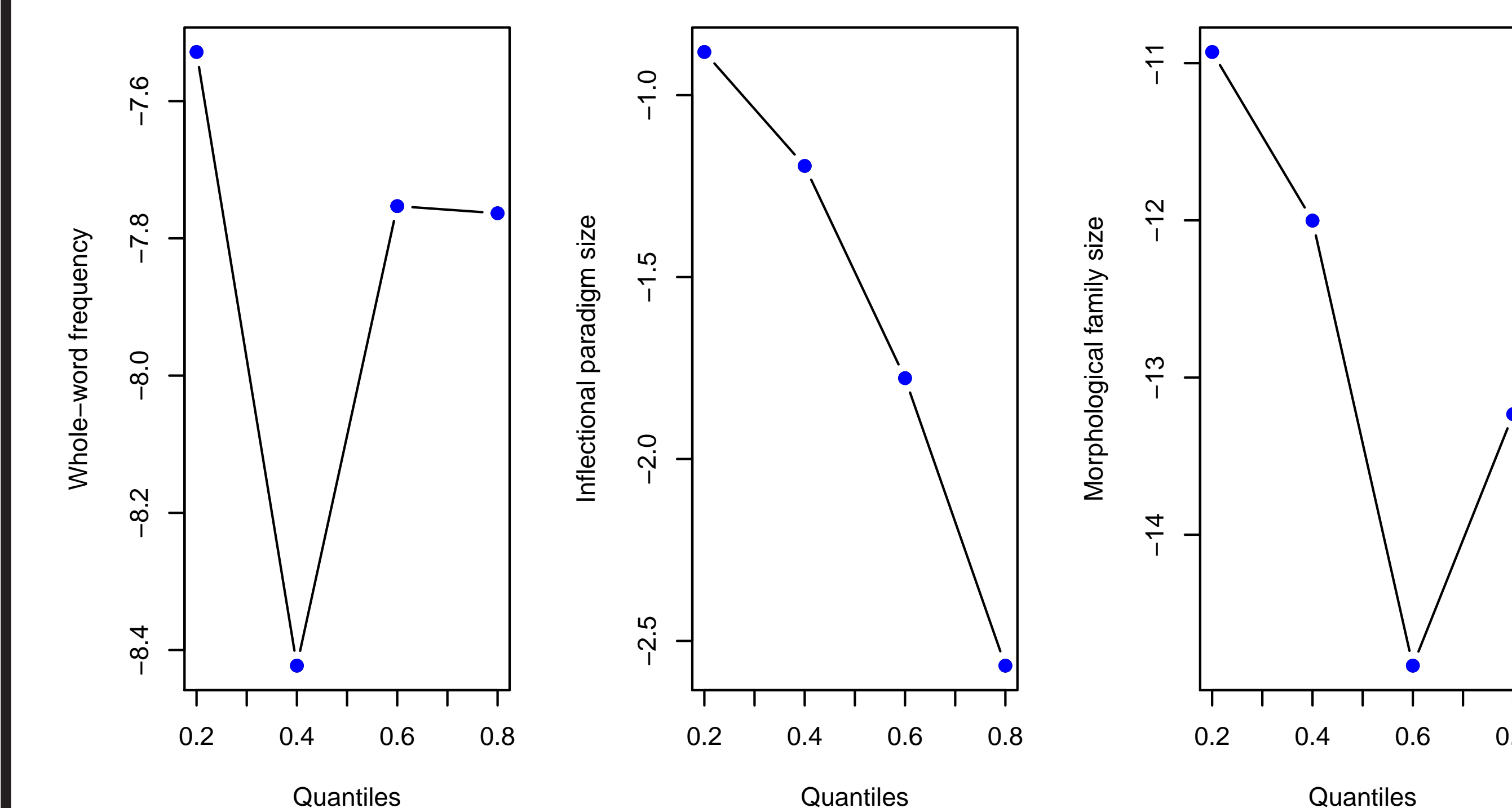

Figure 3: Partial effects for whole-word frequency, inflectional paradigm size morphological family size in the second, fourth, sixth and eighth decile of production latency in Experiment 2.

RESULTS

- **Experiment 1**: production latencies decreased linearly with increasing morphological family size ($\hat{\beta} = -0.0041, t(4721) = -5.1651, p < 0.0001$); acoustic durations with increasing whole-word frequency ($\hat{\beta} = -0.0114, z(4719) = -4.2573, p < 0.0001$; Fig. 1).
- **Experiment 2**: production latencies decreased with increasing whole-word frequency ($\hat{\beta} = -0.0114, z(12026) = -6.6509, p < 0.0001$, morphological family size ($\hat{\beta} = -0.0042, z(12026) = -3.5101, p = 0.0004$), and inflectional paradigm size ($\chi^2 = 18.4492, edf = 3.0310, p = 0.0008$); the same effect emerged in acoustic durations (Fig. 2).
- **Quantile regression**: whole-word frequency peaks at the .4 decile; inflectional paradigm size and morphological family size peaks later on at the .8 decile and .6 decile (Fig. 3).

DISCUSSION

- The whole-word frequency effect for inflected forms is present for all inflected nouns words, not only irregular ones; the effect is the strongest earlier in time.
- A novel paradigmatic measure, inflectional paradigm size, facilitates word production; the effect is the strongest later in time.
- The effect of morphological family size was present in the production of case-inflected nouns; the strongest later in time.

CONCLUSIONS

- Whole-word frequency effect for Estonian case-inflected nouns is in line with frequency effects for sequences of words in English (*into the house*, see e.g., Arnon & Snider 2010).
- The emergence of inflectional paradigm size fits well with the well-established effect of morphological family size.
- The amount of information about word use in the mental lexicon seems to be substantially larger than previously assumed.

REFERENCES

Arnon, I. & N. Snider (2010). Syntactic probabilities affect pronunciation variation in spontaneous speech. *JML* 62, 67–82.

Baayen, R. H., T. Dijkstra & R. Schreuder (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *JML* 36, 94–117.

De Jong, N. H., L. B. Feldman, R. Schreuder, M. Pastizzo & R. H. Baayen (2002). The processing and representation of Dutch and English compounds. *Brain and Language* 81, 555–567.

Fasiolo, M., G. Y., R. Nedellec & S. N. Wood (2016). *Fast calibrated additive quantile regression*. R package version 1.0.

Levelt, W. J. M., A. Roelofs & A. S. Meyer (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22, 1–38.

Milin, P., V. Kuperman, A. Kostić & R. Baayen (2009). Paradigms bit by bit. In J. P. Blevins & J. Blevins (eds.), *Analogy in grammar: form and acquisition*, Oxford, pp. 214–252.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language*.

Wood, S. N. (2006). *Generalized Additive Models*. New York: Chapman & Hall/CRC.

ACKNOWLEDGEMENTS