ESUKA – JEFUL 2025, 16–2: 217–256

STABILITY OF MEANINGS IN BASIC VOCABULARY LISTS OF URALIC LANGUAGES

Mervi de Heer[a], Luke Maurits[b]
[a] *Uppsala University, SE*
[b] *Max Planck Institute for Evolutionary Anthropology, DE*
mervi.deheer@moderna.uu.se, luke_maurits@eva.mpg.de

Abstract. Meanings in basic vocabulary lists are expected to be stable. In the literature, various basic vocabulary lists have been proposed that are founded on various criteria, with stability being a key criterion. However, the stability of basic meanings and to what degree popular lists fit these is rarely assessed for whole language families. In this paper, we define stability as resistance to borrowing, and we examine quantitatively how prone to borrowing the meanings in popular basic vocabulary lists are for the Uralic language family. We refine the methodology for meaning-ranking by calculating a borrowing probability for each meaning and clustering them into groups of more borrowable and less borrowable meanings. Our quantitative analysis provides a higher resolution within the basic vocabulary. While basic vocabulary lists are a good fit for the Uralic family in general, we demonstrate that not all basic meanings are equally stable or necessarily basic in the strictest sense.

Keywords: basic vocabulary, Swadesh list, meaning lists, loanwords, borrowing, quantitative methods

DOI: https://doi.org/10.12697/jeful.2025.16.2.07

1. Introduction

The vocabulary of a language can be categorized into basic vocabulary and cultural vocabulary. The latter is understood to consist of words denoting innovative concepts, artifacts, and institutions such as religious terminology or names of agricultural products and techniques. In contrast, the definition of basic vocabulary is less straightfoward. Basic vocabulary – a notion which was acknowledged already in the seventeenth century (Hymes 1983: 65) – is often presented in the literature as semantic groups of general concepts that are essential for human life and communication (Swadesh 1952: 457; Millar & Trask 2015: 21).

These concepts are thought to be expressed in most languages of the world and they include e.g. body parts, kinship terms, natural features, personal pronouns, low numbers, and simple actions (Campbell & Poser 2003: 263; Swadesh 1955: 124–125). The words denoting such basic meanings tend to be frequent, morphologically simple, and semantically unmarked. Cultural vocabulary items in turn often have the opposite characteristics and it is hypothesized that they are often borrowed due to some need or to fill gaps in a system (Matras 2009: 50; Haspelmath & Tadmor 2009: 46).

One of the most important attributes differentiating the categories of cultural and basic vocabulary is the stability of the words that fill vocabulary lists' meaning slots, especially regarding resistance to borrowing. It is expected that basic vocabulary is borrowed more rarely and therefore more accurately reflects inheritance, in other words, vertical relationships between languages, while cultural items tend to show horizontal relationships (for example borrowing) across languages. For this reason, basic vocabulary lists are commonly used in historical-comparative linguistics and in computational applications thereof to investigate ancestral stages of languages (e.g. Campbell 1999: 12; Rankin 2003: 187–188; Embleton 1986; Syrjänen et al. 2013; Heggarty et al. 2023). A strength of standardized basic vocabulary lists is that they are highly useful for crosslinguistic research.

There is no consensus as to which concepts are the most basic as, in principle, none are universally resistant to borrowing (Fox 1995: 95). In contact situations involving deep bilingualism and strong cultural pressure, borrowings can be acquired even within the basic vocabulary (Field 2002: 153; Thomason & Kaufman 1988: 14). Borrowing may play a notable role in the composition of basic vocabulary, because it is unlikely that all meanings from the supposed basic categories are equally stable.

Due to the fuzziness of the notion of basic meanings, extant basic vocabulary lists are of different sizes. What these lists have in common is that they attempt to standardize a selection of the most borrowing-resistant meanings. Standardized basic vocabulary lists are closely linked with the work of Morris Swadesh (1952, 1955), whose lists in 200-item and 100-item versions have become popular. Other lists are often derived from and compared with the Swadesh lists. In these early lists from the 1950s, Swadesh selected the meanings rather intuitively,

led by his scholarly expertise, and formal criteria for meaning selection were not explicitly laid out. However, he stated that basic vocabulary lists should contain universal meanings and refer vaguely to "things found everywhere in the world" (Swadesh 1952: 457). We interpret that his description implicitly contains attributes now commonly associated with basic vocabulary, namely frequency and semantic neutrality, which have since been explicitly considered – among other criteria – in attempts to define basic vocabulary more precisely (e.g. Tadmor 2009: 68). Common to many of the newer lists is that they have their roots in the Swadesh lists and aim to improve on them. Items in the Swadesh lists are usually presented in alphabetical order and are not ranked according to stability (see Holman et al. 2008 and Starostin 2010 for ranked versions involving several language families and lexicostatistical methodology).

Generally, two approaches for list optimization have been pursued. The first one involves the refinement of item-scoring methodologies for a maximally universal list that can be applied to most languages (e.g. Starostin 1998–2005; Dolgopolsky 1986). Here, the goal is to establish a general ranking of meanings from the most borrowing-resistant to the least resistant. The most borrowing-resistant meanings form the core basic vocabulary within the whole possible basic vocabulary of a language. This approach may result in a very low number of meanings (e.g. Holman et al. 2008 presents a list of 40 meanings). The Word Loanword Typology Project (WOLD = World Loanword Database) (Haspelmath & Tadmor 2009), described below as an example, set out to improve the item-scoring methodology. The vocabulary lists were developed using empirical data in order to compile a selection of borrowing-resistant meanings for crosslinguistic research and are applied in the present study.

The second approach is the supplementation of a standard list with items relevant to the environmental context of a particular language group (e.g. Matisoff 1978). This aims to address the problem of missing data since the meanings in general lists are based on broadscale approximations. Language family-specific meanings are left out or not all meanings are equally relevant (McMahon & McMahon 2005: 41).

Some basic vocabulary lists have been developed by combining both approaches (e.g. McMahon et al. 2005), while others have sought greater accuracy by heuristically dropping all irrelevant meanings.

This can cause gaps in the data without suppletion from a readily available list (e.g. Bastin et al. 1999). List (2023) summarizes current and past repositories of crosslinguistic wordlists which include basic vocabularies.

Research is limited on how resistant generalized basic vocabulary lists and their meanings are to borrowing when large, family-wide language samples are examined. To our knowledge, this information is unavailable for most language families, but some prior studies do exist. In a large-scale study of Indo-European basic vocabulary containing the Swadesh200 list but also involving different data and methodology, the average percentage of borrowing is 8% (Nelson-Sathi et al. 2011). Dybo (2013) presents information on semantic changes affecting the acquisition of basic vocabulary and numbers of borrowings for the Swadesh100 meanings of Turkic languages for the purposes of glottochronological analyses. Still, it is not possible, due to limited comparable studies, to give a clear-cut answer as to whether the basic vocabularies of the Uralic language family are more prone to borrowing than in other families.

A fresh take on developing basic vocabulary lists came with the Loanword Typology Project (also published as WOLD = World Loanword Database, Haspelmath & Tadmor 2009). This study and the database produced by its results provide empirical evidence for the borrowability of vocabulary, based on large and thorough surveys consisting of 41 languages representing various language families (Haspelmath & Tadmor 2009: 4–5). The meanings in WOLD (1460 in total) are ranked quantitatively according to four main criteria: susceptibility to borrowing, historical age within the language family, morphological simplicity, and representativeness of meaning in different languages (Haspelmath & Tadmor 2009). The evaluation of the criteria for the language sample has produced the 100-item Leipzig–Jakarta (= LJ) list, which is the most borrowing-resistant selection of the meanings found in WOLD. The LJ list can be considered as an improved alternative to the Swadesh lists, but it has not yet been applied to the same extent. In recent years LJ has been used, for instance, to study relationships between Turkic, Mongolic, Tungusic, Japonic, and Koreanic in seeking evidence for a macrofamily hypothesis (Robbeets 2020). The amount of borrowing found in the LJ lists of Turkic and Mongolic and the low likelihood of a common origin for these language families has been analyzed in Erdal (2019: 96).

For the Uralic language family, the Swadesh lists are the most commonly used basic vocabulary lists (e.g. Sammallahti 1998 for Saamic languages). A research example on a whole subgroup of Uralic is provided by Rozhanskiy & Zhivlov (2019), where Finnic Swadesh100 lists are analyzed and information on borrowings is provided. Uralic basic vocabulary meanings are also included in large datasets such as the NorthEuraLex 0.9 database, where the total list has 1016 meanings (Dellert et al. 2019). Furthermore, the full WOLD survey has been conducted as a part of the WOLD project for Kildin Saami (Rießler 2009) and separately for Finnish (Cronhamn 2018). Still, on the whole, borrowing of basic vocabulary in Uralic languages remains an insufficiently researched topic (cf. De Heer et al. 2024). In this paper, we set out to fill in this gap in knowledge by exploring the resistance to borrowing and susceptibility to borrowing of basic vocabulary meanings for the Uralic language family using the Swadesh lists and vocabulary lists from the WOLD project, the Leipzig–Jakarta, and another selection called WOLD401–500.

The overarching aim of this paper is to examine borrowing resistance in basic vocabulary. More precisely, we study how stable or borrowable the meanings in generalized basic vocabulary lists are from a Uralic perspective; i.e. how well the readily available lists fit the Uralic family. We chose to focus on borrowing resistance because it is a baseline criterion for basic meaning selection. As a secondary goal, we develop meaning-ranking methodologies to factor in historical and geographical circumstances affecting the stability of meanings in Uralic.

The Uralic language family is especially compelling for studying borrowing resistance in basic vocabulary because borrowing information is accessible: for this language family there has been extensive historical-comparative research, producing a vast body of literature on loan etymologies. A novel type of basic vocabulary dataset combining popular basic vocabulary lists called UraLex version 2.0 has been released for Uralic (De Heer et al. 2021). De Heer et al. (2024) presents loanword information for the dataset and a detailed study on the loanword strata in a sample of six languages.

We quantitatively analyze the stability and borrowability of the meanings contained in a Uralic basic vocabulary dataset. We seek to fulfill our main aims by means of three specific tasks:

1. We measure whether the ranked Leipzig–Jakarta meanings are also highly stable in Uralic and evaluate how the Uralic data confirms or deviates from the ranking based on the worldwide sample.
2. We make a prediction that the LJ list has more stable meanings and less borrowing-prone meanings than the Swadesh lists for Uralic. We also predict that the WOLD meanings with ranks 401–500 are more borrowable. We assume this because the ranked lists are based on a crosslinguistic empirical study of borrowability instead of intuitively selected criteria.
3. We identify the meanings that conflict with the expectations of stability and borrowability implied by the standardized lists.

2. Materials and Methods

2.1. The UraLex 2.0 basic vocabulary dataset

We use the published dataset *Uralic basic vocabulary with cognate and loanword information* (version 2.0), hereinafter referred to as UraLex 2.0 (De Heer et al. 2021, see Syrjänen et al. 2018 for version 1.0; Syrjänen et al. 2013 and Lehtinen et al. 2014 on applications of earlier versions on the data).

The dataset is a composite of four basic vocabulary lists and, after overlap between the lists has been taken into account, has 313 meanings in total. The amount of overlapping meanings and the percentage of overlap is given in Table 1. Multiple lists were selected to provide more data for quantitative analyses while still retaining comparability between languages. The data contains the meanings of the Swadesh100, Swadesh200, Swadesh207 (= Swadesh200 and Swadesh100 combined, including 7 meanings present in Swadesh100 but not in Swadesh200), and Leipzig–Jakarta, as well as somewhat more borrowable meanings, the WOLD 401–500 rankings (introduced in Syrjänen et al. 2013). There are minor adaptations, e.g. *foot* and *leg* appear twice in UraLex following the convention of the Swadesh200 and Swadesh207 lists. This means that as the concept *foot/leg* also appears in the LJ list, the total number of LJ meanings is 101. Within UraLex, the more concise lists Swadesh100 and LJ represent a stricter interpretation of basic vocabulary, while the Swadesh207 is a less conservative selection. The WOLD 401–500 meanings are included to increase the amount of data,

but also to test how far the expectation for the borrowing-resistance of basic vocabulary can be extended. These meanings were selected as an expansion because they partially overlap with the Swadesh207 list but also represent other relatively "basic" meanings when compared to the whole WOLD list covering over 1000 meanings. Together, the composite data represent basic vocabulary in the widest sense.

Table 1. Presentation of basic vocabulary lists in UraLex 2.0 and the amount of overlap between the lists. The A column contains the 100-item and 200-item Swadesh lists and their combination as Swadesh207. The B column contains the Leipzig–Jakarta and WOLD401–500 lists. The column "Overlapping meanings and %" shows the raw number for how many meanings are the same between the lists and also the percentage of overlap. There is no overlap between the Leipzig–Jakarta and the WOLD401–500 lists. The total number of meanings in UraLex 2.0 is 313.

List A	List B (n)	Overlapping meanings and %
Swadesh100	Swadesh200	93 (93.0%)
Swadesh100	Leipzig–Jakarta (101)	63 (63.0%)
Swadesh100	WOLD401–500 (100)	3 (3.0%)
Swadesh200	Leipzig–Jakarta (101)	79 (39.5%)
Swadesh200	WOLD401–500 (100)	13 (6.5%)
Swadesh207	Leipzig–Jakarta (101)	82 (39.6%)
Swadesh207	WOLD401–500 (100)	13 (6.3%)

UraLex 2.0 also includes tagging expressing which words occupying the meaning slots share the same historical origin covering both inheritance and borrowing. For example, when one examines the Permic and Finnic words for the meaning 'tooth' (a word that exists both in the LJ and Swadesh lists), it is clear that the Permic words are inherited from Proto-Uralic. The Permic words are placed into one class and designated with the label 'a'. The Finnic words in turn are indisputably borrowings from the Baltic branch of Indo-European and are designated 'b'. Not all reflexes for the meanings are undivided free morphemes, but untangling the history of derivations is a task best suited for later research. Words sharing the same inherited root in the same meaning slot are grouped together even if the words are derivations. Thus, in UraLex 2.0, words are treated as root-meaning traits which capture historical relationships on a wide scale, following the root-meaning approach introduced by

Chang et al. (2015). If a word does not have a historical relationship with a word from another language in the same meaning slot, the word gets its own class.

Another attribute of the UraLex 2.0 dataset is that synonyms are allowed in meaning slots. This means that each language in the dataset has a varying number of words for a basic vocabulary list. Synonymy was permitted in order to achieve maximal representation and coverage of languages.

We study only the stability of the basic vocabulary meanings with regard to borrowing. Therefore, words acquired solely through borrowing affect our borrowability scores. In addition to borrowing, other types of linguistic processes such as semantic drift, innovation through derivation, compounding, and lexicalization have an effect on how words have ended up in the meaning slots. We leave a detailed analysis of the results of such processes in UraLex for future research and only mention them when relevant.

The source languages of the borrowings in the meaning slots were identified using etymological resources (see De Heer et al. 2024 for details) consisting of articles on individual etymologies, etymological dictionaries, and evaluative literature discussing the credibility of the loan etymologies. The evaluative literature and discussion in other etymological literature was used to assign certainty estimates to the borrowings, allowing us to partition the data into sets ranging from a more uncontroversial set with generally accepted borrowings (all borrowings assigned as *clear* borrowings) to a set with all suggested borrowings, including etymologies that still require thorough evaluation (= *clear*, *probable*, and *possible* categories together). In this paper, the focus lies on the set representing a middle point in conservativeness. We consider the loanwords with *clear* and *probable* certainty estimations but leave out the more uncertain borrowings with the *possible* tag.

The most prominent borrowing sources in the data and the representativeness of the loanword layers is discussed in De Heer et al. (2024). The study reveals that the basic vocabulary of Uralic languages contains both loanwords acquired at the ancestral stages of the Uralic family as well as later borrowings into individual languages. The certainties of loanwords vary such that newer independent borrowings tend to be uncontroversial and explicitly tagged as *clear*. Older borrowings, on

the other hand, are often fraught with considerable uncertainty, which is addressed by tagging them as *probable* or *possible*.

UraLex 2.0 comprises 25 existing languages and an approximation of reconstructed Proto-Uralic. All languages present in UraLex 2.0 were included in our analysis. Of the languages, 12 languages have borrowing information meeting our criteria of conservativeness (loanwords for a meaning having a certainty estimation of *clear* or *probable*) which can directly inform our analysis. The total number of basic vocabulary meanings is 313 and all meanings are included in the analyses. In Sections 2.4 and 3.2, only the meanings from the LJ list are examined. From the total number of meanings (313), a set of 198 meanings have at least one borrowed reflex in a language. The information provided by them is used to estimate how borrowing-prone basic vocabulary meanings are for the Uralic language family using Bayesian logistic regression (see Section 2.2 for details). The method allows one to answer binary questions, i.e. we could formulate a question to ask of our analysis, "Is a basic vocabulary meaning borrowing prone? Yes/No." The Bayesian method incorporates prior knowledge (here borrowing information, subgroup, and geographical information) to guide the answer. We call the results "borrowability scores", which are values representing the likelihood of a meaning being borrowing-prone given the priors and our data. This has been done for all meanings using all languages.

Next, as laid out in Section 2.3, the rather homogeneous data (313 meanings with "borrowability scores" still being relatively close to each other) is then divided into clusters with the help of the "borrowability scores" to identify the most extreme cases of borrowable meanings and those that are most likely not borrowed. A threshold is calculated to aid the analysis. The scores, probability distributions, and clusters are visualized in Figure 1. The probabilities for some meaning being borrowed are visualized as violin plots. Here, the thickness of a violin tells how likely a meaning is borrowable or stable, i.e. the thicker the violin, the higher the likelihood. In other words, the violins show a distribution of probability densities. To form clusters, the whole bulk of 313 meanings is needed for resolution in our type of data. The clusters are used to examine the differences between the stability and borrowing-proneness of meanings and the properties of the basic vocabulary lists as a whole. A detailed description of all methods with technical details is given below. An introduction to Bayes' theorem and

examples of incorporation of prior information for analyzing linguistic data is available in, for example, Greenhill et al. (2020: 229–230). While the focus of those authors is on phylogenetics, the general introduction to Bayesian methods applies to this paper as well.

2.2. Measuring meaning stability and borrowability

Several aspects of the data and our research tasks hinder a straightforward application of familiar statistical tests for comparing proportions across different samples. Many statistical tests that are common within the paradigm of frequentist method testing, involve basic assumptions regarding the data. These include, for example, randomness, independence of observations, and that the data be normally distributed. An introduction to common frequentist methods applied to linguistic data is given in, for example, Woods et al. (1986). For basic vocabulary data, one complicating factor is that the different lists are not well modeled as small random samples from distinct large statistical populations, i.e. every member in a group. For instance, such populations could be tokens in a linguistic corpus or all pupils attending a certain school. The meanings in the lists are not chosen randomly but because of their (seeming) resistance to borrowing, and moreover the same meaning can appear in more than one list. Another complicating factor is that the observations of whether or not a particular word form has been borrowed are not independent due to a number of reasons, e.g. borrowing into a common protolanguage or overlap of lists.

For these reasons (non-random samples and dependencies), we approach the problem of comparing lists indirectly by estimating a "borrowability score" for each of the 313 meanings independently of their presence or absence from any particular vocabulary list. The lists can then be compared straightforwardly by looking at summarized statistics regarding the distributions of borrowability scores over different lists.

It may seem that the available data has a limited ability to address the questions of interest, as for each of our 313 meanings we have only a count of the number of languages with a borrowed form for that meaning, and this count is necessarily between zero and 12 (the number of languages in UraLex 2.0 for which borrowing information is available). It is not immediately obvious how such data can do more than

sort the meanings into at most 13 different categories of presumably equal borrowability, resulting in quite a low-resolution ranking of 313 meanings. However, we can do better than this by considering the genealogical relatedness and spatial proximity of our languages.

Suppose the forms for meaning A occur as borrowings in Finnish, Estonian, and North Saami, while the forms for another meaning B occur in Finnish, Udmurt, and Selkup. Though each meaning occurs as a borrowing in three languages in our data, it would be misleading to assume that meanings A and B are approximately equally borrowable. The fact that Finnish and Estonian are very closely related, together with the geographic proximity of the Finnic and Saamic languages, make it plausible that the three observations of borrowed forms correspond to a single actual act of borrowing, which subsequently was inherited. Such an explanation is substantially less plausible for an equal number of more genealogically and geographically diverse languages, and therefore the data provide stronger evidence that meaning B is more readily borrowed than meaning A. Thus, we can estimate the relative borrowability of meanings even if superficially we have seemingly equivalent data for them.

Nevertheless, two things remain true. One is that we cannot easily say anything about the relative borrowability of those meanings for which we have no borrowed forms in our Uralic data. Some 115 of our meanings fall into this category, and thus we must be content with a ranking of the remaining 198 meanings which do have at least one borrowed reflex in the languages that UraLex comprises. The other fact is that 12 languages with adequate borrowing information can necessarily provide only a limited amount of information, as does including considerations of genealogy and geography. We must therefore expect a considerable degree of uncertainty in our rankings. Highly generalized hypotheses and typologies (see Matras 2009: 156–157) on borrowability based on e.g. parts of speech provide valuable context for interpreting our results, though our analysis draws from concrete data.

In light of this unavoidable uncertainty, we have opted to perform our analysis using Bayesian statistics. In contrast to frequentist or "classical" statistical methods, Bayesian methods provide explicit information on the uncertainty which remains in estimates of model parameters after taking the data into account. We perform a Bayesian logistic regression with a single fixed effect of word meaning, plus random intercepts of

language, the Uralic subfamily, and a broad geographic categorization of the languages into "East", "Middle", and "West". In essence, this model means that each individual meaning receives its own probability of having a borrowed form in a particular language, but that borrowing probabilities for all meanings can be simultaneously raised or lowered to some degree in specific languages, subfamilies, or regions. A random intercept helps to model variability in the baseline probability for a meaning being borrowed when taking informative groupings (geographic category, language, subfamily) into account.

The rough three-way categorization is based on the current areas in which the Uralic languages are spoken, uses natural formations as guidelines, and takes a horizontal point of view. The exception to this view is Hungarian, which we placed in the East group despite its speakers' eventual migration to the Carpathian Basin. The East category contains the UraLex 2.0 languages belonging to the Samoyedic and Ob-Ugric subgroups separated from the Middle group by the Ural Mountains. The Middle group in proximity to the Volga River area contains Meadow Mari from the Mari group, Erzya from the Mordvinic languages, and the Permic languages. Under the West group are the subgroups Finnic and Saamic, which are spoken in Fennoscandia or – in the case of the eastern Finnic languages – relatively nearby in westernmost Russia. We also include the southernmost Finnic languages under the West group. The geographical grouping is largely mirrored by genealogical grouping involving loanwords and phonological evidence. For example, the Finnic and Saami subgroups are tied together in a Western Uralic entity by shared loanwords, phonological evidence connects the languages from the Mari and Permic subgroups to a middle entity, and Hungarian to the east with the Ob-Ugric and Samoyedic languages (Häkkinen 2012: 8–9). In the context of borrowing information, the Mordvinic language present in our data, Erzya, was placed in the geographical grouping Middle. It is likely that Mordvinic has acquired loanwords independently from Finnic and Saamic (Grünthal 2012: 297).

This approach compensates, to some extent, for the effects of inheritance and areal diffusion. Suppose one Uralic subgroup has been in relatively intense contact with one or more other languages for a long time and has consequently accumulated many borrowings. Our model will respond to this by assigning a high random intercept to that subgroup. A meaning whose only borrowings belong to that subgroup will thus not

need to have its own borrowing probability raised as high as a meaning whose only borrowing is in another more typical subgroup. Similarly, if one geographic region has been more isolated from other languages and has fewer borrowings than the other regions, that region will be assigned a lower random intercept. A meaning whose only borrowings are in that region will then have to have its borrowing probability raised quite high to overcome this. This method is far from perfect, and it is still prone to overestimating the borrowability of meanings where only a single borrowing event early in Uralic history has led to borrowed forms being inherited into all or most of the family. Nevertheless, it represents a substantial improvement over previous methods for estimating borrowability, which have not taken such factors into account at all.

The formal definition of this model is as follows, where x_{ij} is a variable indicating whether or not the word for meaning slot i in language j (belonging to subgroup s_j and geographic area g_j) has been borrowed:

$$x_{ij} \sim \text{Bernoulli}(p_{ij})$$
$$p_{ij} = \text{logit}^{-1}(b_i + u_j + v_{sj} + w_{gj})$$
$$b_i \sim N(0, 1.5),\ i = 1,...,313 \qquad \text{(fixed effects of meaning)}$$
$$u_j \sim N(0, \sigma_1),\ j = 1,...,12 \qquad \text{(random effects of language)}$$
$$v_{sj} \sim N(0, \sigma_2),\ s_j = 1,..., 7 \qquad \text{(random effects of subgroup)}$$
$$w_{gj} \sim N(0, \sigma_3),\ g_j = 1,..., 3 \qquad \text{(random effects of area)}$$
$$\sigma_{1,2,3} \sim \exp(1) \qquad \text{(random effect prior)}$$

The final output of the model for our purposes is a set of posterior distributions, one for each meaning, over the relevant meaning's "borrowability score" b_i.

2.3. Clustering the basic vocabulary meanings for comparison

Due to the relatively small number of languages in the dataset, our model returns uncertain estimates of each meaning's borrowability score, and for the majority of pairs of meanings, we are unable to make any kind of definitive statement of the type "our data strongly suggests that meaning A is more/less borrowable than meaning B". Even if one meaning's posterior mean borrowability score is higher/lower than

the others, the 95% HPD (= highest posterior density) intervals of the borrowability scores will most often have substantial overlap. However, we *are* able to make such statements about *some* pairs, particularly when comparing meanings whose borrowability scores are at opposite extremes of the scale. In order not to lose sight of this limited resolution provided by our data, we consider our main result to be not a complete ranking of all 313 meanings from most to least borrowable, but instead a partitioning of the meanings into three clusters.

For any given threshold value of borrowability score, we can divide the meanings into three non-overlapping sets: those whose 95% HPD borrowability score interval is entirely below the threshold (i.e. those meanings which are "almost certainly less borrowable" than the threshold), those whose 95% HPD interval is entirely above the threshold (i.e. those which are "almost certainly more borrowable" than the threshold), and finally those whose 95% interval includes the threshold. We thus seek the threshold value which minimizes the difference in size of the first two sets, i.e. we seek a threshold where we can say with confidence that there are roughly as many meanings less borrowable than the threshold as there are more borrowable. We interpret these resulting and roughly equally sized sets as being the most stable and most borrowable meanings in the dataset, while the third set (containing those meanings which were not clearly more or less borrowable than the threshold) are those for which we can make no strong conclusions.

This approach divides the UraLex meanings into three clusters of approximately equal size: Cluster A (containing 115 meanings) which are the most resistant to borrowing, cluster B (106 meanings) about which we cannot make strong conclusions, and cluster C (92 meanings) which are the most borrowable. This clustering is used to carry out our first research task on the stability of the LJ list for Uralic, and our second task on the comparison of the basic vocabulary lists covered by our dataset.

2.4. Evaluating the stability of the Leipzig–Jakarta meanings for Uralic

We evaluate whether the ranked Leipzig–Jakarta meanings representing the most stable meanings in the WOLD survey also appear as highly stable meanings for Uralic. Here we take advantage of the

borrowability scores of the LJ meanings and the threshold values for clusters A–C. We assess whether the ranked LJ meanings for which we have the borrowability scores from Uralic follow a trend in their clustering: in an ideal scenario, the lowest LJ ranks starting from rank 1 should appear in cluster A while higher ranks should cluster into B and C. We conduct our assessment with the help of a box-and-whiskers visualization of the clustering (see Figure 2). The whiskers represent the minimum and maximum LJ ranks from 1 to 100. The box itself represents the upper and lower quartiles within a cluster (A, B, or C) and how the ranks of the LJ meanings are spread. The boxes further depict in which cluster the ranked LJ meanings are concentrated according to our borrowability scores. The vertical line represents the median rank for the particular cluster. To confirm the visually assessed trend, we measure the similarity of the rank order between the general and Uralic meanings, for which we use Kendall's tau correlation coefficient for ranks. The τ value is calculated by counting the number of pairs of meanings with the same relative borrowability ranking in two lists and subtracting the number of pairs of meanings with conflicting relative borrowability. The difference between these counts is then divided by the total number of pairs of meanings to ensure that values lie between –1 (when the ranking order is perfectly inverted) and 1 (when the ranking order is identical).

We calculate multiple values of τ, corresponding to multiple versions of both rankings. Regarding the LJ list, it sometimes occurs that different meanings are assigned identical borrowability scores by WOLD. For example, the meanings 'name' and 'louse' have the same ranking of 23 (see Tadmor 2009: 69–71 for all ranks and scores for the LJ list). Regarding UraLex 2.0, our Bayesian analysis returns 4000 borrowability values for each meaning (i.e. the posterior samples), reflecting uncertainty as to their true borrowability. For each of the 4000 corresponding rankings, we use a different random tiebreaking of the ranking by WOLD scores, and calculate the resulting τ. Here, instead of treating ties as outliers, random ranks are assigned to the values so that the tie can be resolved. The process helps to weigh a large number of ranking scenarios and avoid bias in the data. This yields a distribution of correlation coefficients which we summarize with the mean value.

2.5. Comparing the basic vocabulary lists

We compare the stability and borrowing-proneness of the whole of the basic vocabulary lists in UraLex 2.0. We examine which list has the largest proportion of its meanings in cluster A of the least borrowable meanings, in the middle ground of cluster B, and in cluster C containing the most often borrowed meanings. The meanings were clustered as described in Section 2.3.

We evaluate whether there is a difference between the accuracy of the data-driven meaning lists LJ, WOLD 401–500, and the qualitative selection of meanings in the Swadesh100 and Swadesh207. Here we define that an accurate basic vocabulary list has the lowest number of borrowing-prone meanings in comparison to other lists and a high number of stable (most borrowing-resistant) meanings. Our comparison allows us to assess how the readily available basic vocabulary lists suit the Uralic family in general.

2.6. Identifying noncompliant meanings in Uralic

Our third task focuses on the meanings expected to be highly stable based on their WOLD ranks but are unstable based on their borrowability scores and vice versa (i.e. meanings expected to be unstable but are found to be highly stable). We discuss the respective clustering of the meanings and analyze the patterns mediated by these conflicting expectations and results qualitatively.

We use occurrence in the list as a guideline to establish borders within the clusters A–C established as described in Section 2.3. This helps us to identify the exact meanings within UraLex 2.0 which do not follow the expectations of stability and borrowability posed by their occurrence in the generalized lists. We expect that the meanings which have survived the strictest curation process in the three basic vocabulary lists (included in the Swadesh lists and LJ) would resist borrowing the most. The WOLD 401–500 meanings falling outside the core lists can be expected to be the most borrowable. This leaves us with 150 meanings for which we have certain expectations regarding their stability and borrowability: 62 stable meanings and 88 more borrowable meanings. We had no particular expectations for the rest; these include the rest of the Swadesh meanings as well as the ranked meanings unique

to the LJ list. We allow the meanings without a prior expectation of their stability to be assigned to a cluster based on their borrowability scores and the threshold values for the clusters.

3. Results

3.1. Overall borrowability distribution

Before comparing our results to previous basic vocabulary lists and stability rankings, we present an example of an overall distribution of borrowability scores for a random selection of meanings from UraLex 2.0. Figure 1 shows the varying levels of borrowability for the meanings and the resolution the clusters give for the data.

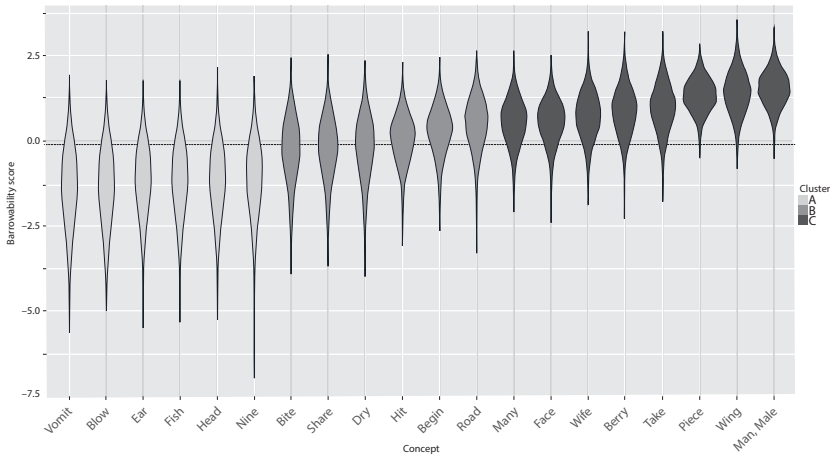

Figure 1. The posterior distribution for the borrowability scores visualized as violin plots with 20 randomly selected meanings from UraLex 2.0. The dotted horizontal line represents the threshold. Meanings extending under the threshold belong to cluster A whereas cluster C is the other way around. In cluster B, the meanings have roughly as much mass on either side.

3.2. The stability of the Leipzig–Jakarta meanings in Uralic

We evaluate whether the ranked LJ meanings in Uralic data follow the expectations set by the general LJ list. The LJ meanings are clustered in Figure 2 according to their borrowability scores and threshold values. Cluster A represents a selection of the least borrowable meanings

within the conservative LJ list whereas clusters B and C contain more borrowing-prone meanings.

The comparison of the clusters reveals that a higher number of meanings assumed to be highly borrowing-resistant (low LJ ranks) group, according to their borrowability scores, into cluster A rather than into clusters B and C. The first quartile of the most borrowing-resistant cluster A contains the largest number of meanings expected to be very stable (LJ rank 25 and below). The median rank in cluster A is much lower (rank 40) than in clusters B and C (rank 55). Clearly higher LJ ranks (above rank 75), which are more borrowable than lower ranks in the generalized LJ basic vocabulary list, are also more borrowable for Uralic as indicated by the third quartile of cluster C.

While the borrowability scores for the Uralic LJ meanings show a trend of low ranks being more stable and higher ranks being less stable, the result is not clear-cut. The rather large clusters B and C show that the Uralic LJ meanings do not strictly follow the general LJ ranking. The first quartiles of both clusters still contain meanings with low LJ ranks in addition to high ones.

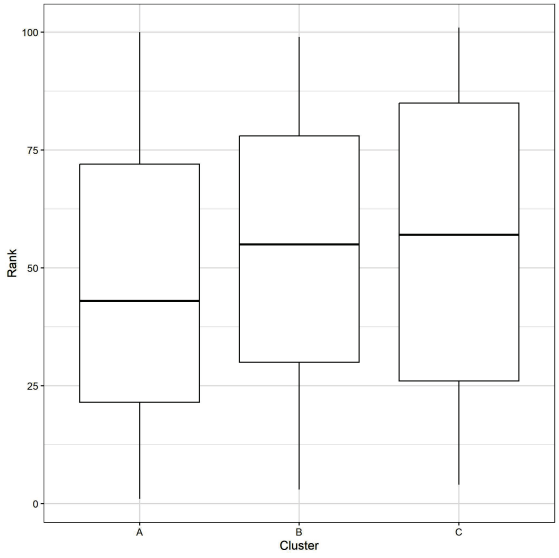

Figure 2. Clustering of the Leipzig–Jakarta meanings for Uralic. The *x*-axis shows the clusters from the most borrowing-resistant meanings (cluster A) to the less-resolved group (cluster B) and more borrowing prone grouping (cluster C). The *y*-axis presents the LJ ranks (1–100).

Next, we investigate whether the LJ meanings appear in a similar order when ranked according to their stability in Uralic. The mean value of the τ coefficient over all our compared rankings was 0.047, with 87% of the values being above 0 (Figure 3). This constitutes clear evidence of a small positive correlation between the rankings.

Only a few cases have a highly divergent rank indicated by the 13% of values below 0. It would not be realistic to assume a perfect agreement in ranking close to τ = 1, as the LJ ranking itself is an approximation of the least borrowable meanings in the world's languages. A τ of 0 would mean that there is no relationship between the two rankings at all.

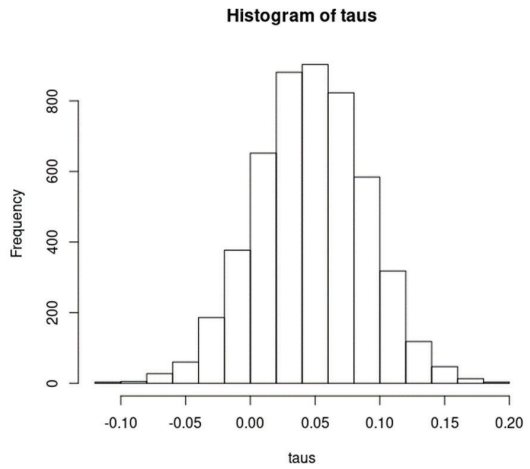

Figure 3. A histogram of taus for the Leipzig–Jakarta meanings in Uralic. The *x*-axis shows that most meanings appear in bins with positive values above zero. The *y*-axis shows how frequent a bin is.

3.3. Comparing the stability of Uralic basic vocabulary lists

We compare the borrowing resistance of all basic vocabulary lists in UraLex 2.0 by examining what proportion of the meanings in each list occur in clusters A–C. We clustered the 313 meanings with the threshold values using the highest posterior density scores between pairs of the most stable and most borrowing-prone meanings in the data. We note that as the size of a basic vocabulary list increases, the proportion of the N most stable words that are included necessarily increases. As Swadesh207 is the longest list in UraLex 2.0, it naturally has more room

for meanings from any cluster. Examining the percentages helps to compare the total lists side-by-side (Table 2).

Table 2. N of items in Cluster A (the most stable meanings), the middle-ground Cluster B, and the more borrowing-prone Cluster C. The percentages indicate how large the proportion of items from cluster A–C is in each list. The basic vocabulary lists' names are shown in the List column.

List	Cluster A (n = 115)	Cluster B (n = 106)	Cluster C (n = 92)
LJ	42% (n = 43)	34% (n = 34)	24% (n = 24)
Swadesh100	45% (n = 45)	34% (n = 34)	21% (n = 21)
Swadesh207	41% (n = 84)	32% (n = 66)	26% (n = 54)
WOLD 401–500	29% (n = 29)	38% (n = 38)	33% (n = 33)

Cluster A contains 115 of the 313 meanings in UraLex 2.0. The meanings in cluster A do not show evidence of borrowing; they are the most stable meanings in UraLex 2.0. They are easily identified by our analysis and are clearly more resistant to borrowing than the threshold. Due to overlap between the basic vocabulary lists, meanings can belong to several lists. Thirty meanings from cluster A are shared by Leipzig–Jakarta and both of the Swadesh lists.

Of all the most stable meanings, 43 are included in the LJ list, 45 in the Swadesh100 list, 84 in Swadesh207, and 29 in the less basic WOLD 401–500 list. While the Swadesh207 list has the largest number of stable meanings from cluster A, they represent 41% of the meanings in the whole list. In turn, 45% of the Swadesh100 list is composed of stable meanings. For LJ, 42% of the meanings come from cluster A. The WOLD 401–500 list contains only 29% stable meanings.

Cluster B covers 106 meanings. It cannot be clearly determined whether they are particularly stable or borrowing-prone within basic vocabulary. Twenty meanings included in all lists, except WOLD 401–500, have been clustered into group B. The short lists LJ and Swadesh100 both have 34 of their meanings from cluster B and Swadesh207 has 66. For the less basic WOLD 401–500 list, 38 meanings come from cluster B. Percentage-wise, cluster B meanings occupy similar proportions of the basic vocabulary lists: 34% of the LJ list as well as Swadesh100. Swadesh207 has a slightly lower proportion: 32%. Of WOLD 401–500, the percentage is 38%.

Cluster C represents the most borrowing-prone meanings in UraLex 2.0 and has 92 meanings. Of these, 12 are shared by LJ and the two Swadesh lists. Leipzig–Jakarta contains 24 more borrowing-prone meanings, Swadesh100 has 21, and Swadesh207 has 54. In total 33 more borrowing-prone meanings appear in WOLD 401–500. The WOLD 401–500 has the highest percentage of borrowing-prone meanings at 33%, while Swadesh207 has 26% and LJ 24%. Finally, 21% of the Swadesh100 meanings come from cluster C.

We expected that the LJ list would contain the most borrowing-resistant meanings for Uralic as the stability of the LJ list is backed up by crosslinguistic sampling. We also expected that WOLD 401–500 would have the largest proportion of more borrowing-prone meanings, because these meanings are borrowed more often according to WOLD. However, examining our clusters reveals that it is the Swadesh100 list that has the largest proportion of its meanings from cluster A. Swadesh100 also has 38 meanings which are not included in LJ. Of those, 15 appear in cluster A. Likewise, 13 meanings unique to LJ are among the most stable meanings. This explains the difference between the shorter basic vocabulary lists. While our expectation does not hold, the difference between Swadesh100 (45% of its meanings from cluster A, i.e. expected to be stable) and LJ (42%) is quite small. The longer Swadesh207 also comes close with 41%. When also taking cluster C into account, the Swadesh100 is the most accurate basic vocabulary list because it has the lowest percentage of its meanings from the most borrowing-prone meanings.

The proportion of WOLD 401–500 meanings from the most borrowing-resistant cluster A is noticeably smaller than in the case of LJ and the Swadesh lists. This confirms that most of the crosslinguistically less stable meanings are also more borrowable in Uralic. The result fulfills our expectation that the core basic vocabulary lists (= LJ, SW100, SW207) are more resistant to borrowing in the case of Uralic, while WOLD 401–500 is the least accurate. This is further confirmed by the fact that almost half (30/62) of the meanings that are shared by the core lists appear in cluster A.

3.4. The noncompliant meanings in Uralic

We identified the meanings in the Uralic data that do not follow the general trends of stability and borrowability implied by the standardized basic vocabulary lists. For resolution, we clustered the meanings according to their borrowability scores, threshold values, and presence in the different basic vocabulary lists included in UraLex 2.0 (Figure 4). We observe common patterns and the most prevalent word-acquisition strategies mediated by our results, remark on the morphology of the reflexes for the meanings. For a detailed etymological analysis of individual words for meanings in UraLex 2.0, see the references in De Heer et al. (2021).

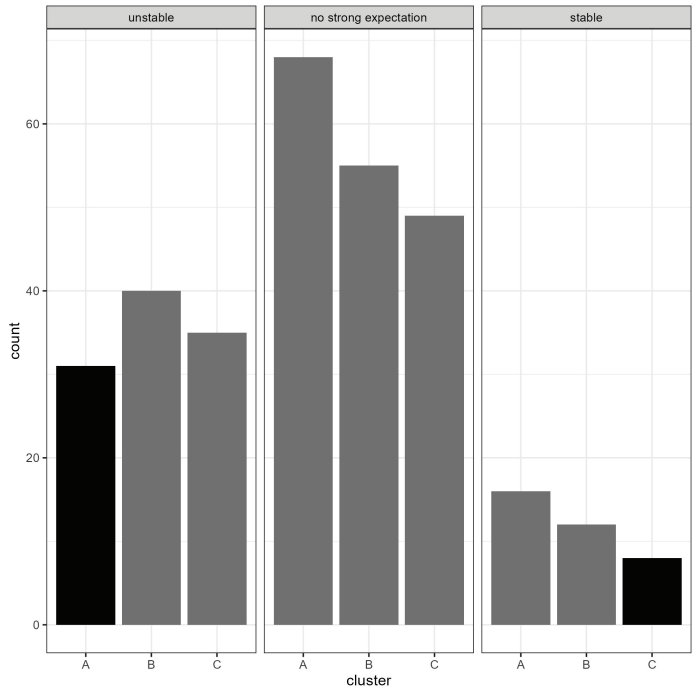

Figure 4. Meanings in UraLex 2.0 clustered based on their borrowability scores and threshold values. The division into panels is done by factoring in the expectations of stability and borrowability. The black cluster A contains the meanings that appear stable while expected to be more borrowable. The black cluster C has meanings appearing more borrowable while expected to be among the least borrowable meanings.

3.4.1. Stable meanings expected to be more borrowable

In total, 25 meanings (Table 3) appear surprisingly borrowing-resistant. They were expected to be more borrowable because they occur in the WOLD 401–500 list without much overlap with the other lists. These words are often inherited and also acquired in multiple innovative ways, but rarely are they borrowed. However, the "unstable" panel in Figure 4 shows that the majority of the WOLD 401–500 meanings tend to be more borrowable in Uralic just as expected.

The reflexes of six meanings ('armpit', 'cooked', 'to divide', 'to extinguish', 'lung', and 'son-in-law') are primarily acquired through inheritance and occur across languages in UraLex 2.0. They represent the clearest examples of meanings that could be expected to be more borrowable based on their higher WOLD rankings. For Uralic they are stable; the meanings are mostly represented by words assumed to be of Proto-Finno-Ugric or even Proto-Uralic origin. A minor word-acquisition strategy for the reflexes of the six meanings is semantic shift of another inherited word, for instance, the reflexes of the meaning 'lung' in the Finnic subgroup may have undergone a change in meaning from 'light' to 'lung' (SSA), which is a common semantic shift (see e.g. Urban 2021). While some languages feature an individual borrowed reflex (for instance, the meanings for 'son-in-law' in Skolt and Kildin Saami are Russian borrowings), the prevalence of inherited words places these meanings among the stable ones.

Common for these meanings is that they reflect true cognate relationships well. The languages have mostly retained both the linguistic material and the original meaning. This means that the number of reflexes acquired through derivation and semantic shift is not very high. The meanings require only a small number of root-meaning classes across the languages present in UraLex 2.0.

For ten meanings ('to cut down', 'to go out', 'to keep', 'to light', 'nine', 'shallow', 'to shit', 'stinking', 'to stretch', 'to touch') the reflexes are mostly derived from inherited stems. The higher number of root-meaning sets the reflexes require show that, besides derivation, also semantic shifts or expansions have likely taken place. For instance, the Permic reflexes for 'to keep' may have developed from the meaning 'to look' (Csúcs 2005: 397).

The reflexes of two meanings (‘to groan’ and ‘to mumble’) have been innovated on an expressive basis independently. Therefore, they are unlikely to reflect deep inherited relationships.

The reflexes for four meanings (‘be hungry’, ‘to dream’, ‘heel’, and ‘molar tooth’) are mostly expressed as analytic forms for the verbs, compounds for the nouns, or derived from inherited stems as a less common strategy. These meanings are weaker candidates for inclusion in basic vocabulary lists from the Uralic perspective, because their reflexes tend to be morphologically complex and the possibility of pattern borrowing needs to be considered.

Three meanings (‘grandson’, ‘to untie’, and ‘younger brother’) are falsely grouped as less borrowable meanings by our borrowability scores. They are missing from our dataset for approximately half of the languages, making them appear similar to the group with high numbers of true cognates. The sparse reflexes present in UraLex 2.0 tend to be borrowed, for example, ‘grandson’ is a Russian borrowing in Komi-Permyak. (Data on this meaning is missing, however, from UraLex 2.0 for most other languages.) It is questionable whether these meanings are suitable for a basic vocabulary data set in the first place.

Table 3. The 25 meanings, arranged alphabetically, that should be more borrowable due to their higher WOLD (401–500) ranking but appear surprisingly stable in Uralic. The part of speech, semantic field, and ranks are from Haspelmath & Tadmor (2009). The Mechanism column describes the primary acquisition strategy of the reflexes for the meanings.

Meaning	Mechanism	Part-of-speech	WOLD semantic field	WOLD rank
armpit	Inheritance	Noun	The body	500
to be hungry	Analytic forms, expressive innovation	Verb	Food and drink	424
cooked	Inheritance	Adjective	Food and drink	414
to cut down	Derivation from inherited stems	Verb	Basic actions and technology	405
to divide	Inheritance	Verb	Spatial relations	490
to dream	Analytic forms, semantic shift	Verb	The body	483
to extinguish	Inheritance	Verb	The physical world	479

Meaning	Mechanism	Part-of-speech	WOLD semantic field	WOLD rank
to go out	Derivation from inherited stems	Verb	Motion	423
grandson	(missing data), borrowing	Noun	Kinship	481
to groan	expressive innovation	Verb	Emotions and values	443
heel	compounding, derivation from inherited stems	Noun	The body	408
to keep	Derivation from inherited stems	Verb	Possession	491
to light	Derivation from inherited stems	Verb	The physical world	410
lung	Inheritance, semantic change	Noun	The body	486
molar tooth	Compounding	Noun	The body	426
to mumble	Expressive innovation	Verb	Speech and language	480
nine	Derivation from inherited stems	Function word	Quantity	413
shallow	Derivation from inherited stems	Adjective	Spatial relations	468
to shit	Derivation from inherited stems	Verb	The body	494
son-in-law	Inheritance	Noun	Kinship	462
stinking	Derivation from inherited stems	Adjective	Sense perception	457
to stretch	Derivation from inherited stems	Verb	Basic actions and technology	428
to touch	Derivation from inherited stems	Verb	Sense perception	425
untie	(missing data), derivation from inherited stems	Verb	Basic actions and technology	403
younger brother	(missing data), compounding, inherited stems	Noun	Kinship	455

3.4.2. More borrowable meanings expected to be stable

Our analysis reveals eight meanings which were expected to be borrowing-resistant but are highly borrowable in Uralic (Table 4). Not every reflex of the meaning necessarily has evidence of borrowing, but it is the most prominent word acquisition strategy in many subgroups. The Uralic family has borrowed words in various contact situations and at different time depths from Indo-European, Turkic, and other Uralic languages. Common for the meanings is that while they were expected to be highly borrowing-resistant due to their inclusion in the more curated Swadesh100 list, they are not a part of Leipzig–Jakarta.

The reflexes of the borrowable meanings were mostly acquired during independent borrowing events into Uralic languages or their immediate predecessors such as Proto-Finnic or Proto-Permic. Words have been borrowed several times across Uralic, at least in the groupings designated "West" and "Middle". Since such a small number (8) of meanings appear surprisingly borrowable, it is more fruitful to focus on the time depth (in the distant past or more recently) and dynamics (variety of donors from other language families or other Uralic languages) of contact instead of looking for semantic patterns. It can be noted that half of the items represent adjectives, a class regarded as borrowed often. According to studies on borrowing hierarchies, nouns are most borrowable while adjectives are less borrowable but still highly placed in the hierarchies (Matras 2009: 157).

For the West grouping, the words for the meanings for the quantifier 'all' include Baltic borrowings in Proto-Finnic, whereas family-internal borrowings have been identified for Saamic languages. Komi (Zyrian and Permyak) representing the Middle group has acquired the word from Russian. The quantifier 'many' is similar as it is borrowed in the West and Middle groupings: Saamic features Finnic loanwords but there are also borrowings from North Germanic. Mari has borrowed from the Bulgar Turkic language spoken in the Volga region.

There are two nouns in the "surprisingly borrowable" group. The meaning 'fat' in UraLex 2.0 has *probable*-status borrowings from Proto-Germanic acquired into the West group in the case of the Finnic and Saamic languages. Both Finnic and Saamic have likely borrowed the words independently. Erzya representing the Middle grouping and Hungarian for East have borrowed the words from Russian and a Slavic

source, respectively. The meaning for ‘seed’ is widely borrowed and a good example of a meaning which could be expected to be stable due to its inclusion in the Swadesh list, but in fact it is frequently borrowed across Uralic. The Saamic and Finnic languages have borrowings from North Germanic and Baltic, which were acquired during their shared past but also independently. However, it is still debated whether Saamic has had independent contacts with Baltic. The Permic languages in the Middle grouping share words from Volga Bulgar Turkic that were acquired into Proto-Permic. In the East grouping, Mansi has borrowed from a Turkic source (Tatar).

The adjective ‘round’ in the West grouping has been borrowed into Finnic from Proto-Germanic, but the Finnic group also shows more recent contact with Russian. For the Middle grouping, Erzya has a Russian borrowing whereas Mari has a Chuvash loanword. ‘Green’ and ‘yellow’ are frequently borrowed in Uralic and reflect the multilayered contacts with ancient and recent donor languages. For ‘green’, Finnic has old Indo-Iranian and Baltic as well as new Russian borrowings. The Middle languages also reflect similarly old contacts with Indo-Iranian and newer Russian contacts. For ‘yellow’, we find Baltic loanwords in the case of Finnic, whereas languages in the Middle and East groupings show Turkic borrowings from various sources, such as from Tatar in the case of Mari. Within color terms, the category of green-yellow is often interwoven and considered more borrowable than other color terms (Matras 2009: 87). That ‘green-yellow’ is highly borrowable for Uralic aligns well with this trend.

Finally, the adjective ‘cold’ has been borrowed into Finnic, Saamic, and Erzya. The borrowing situation differs from those above, namely the reflexes were acquired in a more distant past from a more uncertain Indo-European source, perhaps the so-called North-Western Indo-European (see Koivulehto 2001). This kind of borrowing does not necessarily indicate that meanings like ‘cold’ really are borrowed more easily than others in the conservative LJ and Swadesh100 lists. They might have been borrowed once in the distant past. Here, ‘cold’ rather demonstrates that basic vocabulary can retain not only inherited items but also old borrowings for a long time, and therefore certain North-Western Indo-European loanwords are tagged as *probable* borrowings in UraLex 2.0. The most debated items are tagged as *possible* and were not treated as borrowings in our study.

Table 4. Eight meanings (in alphabetical order) are expected to be highly resistant to borrowing due to their inclusion in the Swadesh100 and Swadesh207 lists but appear as highly borrowable. The part of speech and semantic fields are taken from Haspelmath & Tadmor (2009). The Donor column presents the borrowing source languages as broadscale labels from UraLex 2.0.

Meaning	Donor	Part of speech	WOLD semantic field
all	Baltic, Finnic, Russian	Quantifier	Function word
cold	North-Western Indo-European	Adjective	Sense perception
fat	Germanic, Slavic, Russian	Noun	Food and drink
green	Indo-Iranian, Baltic, Russian	Adjective	Sense perception
many	Finnic, North Germanic, Volga Bulgar Turkic	Quantifier	Function word
round	Germanic, Russian, Chuvash	Adjective	Spatial relations
seed	North Germanic, Baltic, Volga Bulgar Turkic, Tatar	Noun	Agriculture and vegetation
yellow	Baltic, Tatar, West Old Turkic, Turkic	Adjective	Sense perception

4. Discussion

Resistance to borrowing is the most common assumption underlying basic vocabulary lists, but it has not been clear how accurately these lists capture the most stable meanings for the Uralic family. We examined and compared quantitatively the stability and susceptibility to borrowing of four basic vocabulary lists.

We developed further a quantitative methodology for measuring borrowability and stability. Our Bayesian approach allows us to express uncertainty regarding the stability of meanings and to take geographical and genealogical circumstances into account. This is useful because not all meanings are equally borrowable: not all words for the meanings have been borrowed into contemporary Uralic languages or as frequently by all languages.

Two major benefits arise from analyzing the stability of basic vocabulary meanings quantitatively. Firstly, we can evaluate how basic vocabulary lists fit a whole language family instead of just a few languages. Secondly, we have obtained more resolution within basic vocabulary regarding meaning stability, i.e. we have identified clusters of more and

less borrowable basic meanings in relation to all meanings. Finding information on borrowing in basic vocabulary can be a challenge for many language families, because most well-known studies on borrowability (e.g. Thomason & Kaufman 1988; Haspelmath & Tadmor 2009) operate with the dichotomy of more borrowable cultural vocabulary vs. borrowing-resistant basic vocabulary, ignoring the differences within the basic vocabulary. However, the resolution we acquired for Uralic is still not detailed enough to provide a strict ranking of all meanings according to stability. A large number of meanings can be described as "mostly resistant to borrowing" and are hard to analyze individually. Our clusters provide a more realistic viewpoint for comparison by identifying "extremities" within homogeneous data. We also identified the meanings that clearly behave against the expectations posed by the standardized lists.

4.1. Comparing the ranking of the standardized Leipzig–Jakarta list with Uralic data

In the first research task we set out to answer the question of whether the meanings in the ranked Leipzig–Jakarta list appear stable for Uralic in the same way.

The LJ list is a compelling alternative to the Swadesh100 basic vocabulary list as it was compiled on the basis of linguistic data instead of heuristic estimates. The meanings in the LJ are ranked from least borrowable to most borrowable. The ranking can be treated as a prediction for what the most basic meanings in any given language might be. We investigated how well Uralic data aligns with the standardized list. We clustered the LJ meanings into three groups from the most stable meanings to the most borrowing-prone ones and examined which cluster had the most meanings with low ranks from the global list. We assessed the similarity between the general ranking and how the meanings rank for Uralic with the Kendall's tau test.

Our results were not clear-cut, but we do observe a tendency of similarity between the general and Uralic rankings: the lowest LJ ranks tend to be highly resistant to borrowing in Uralic as well. Naturally, the small number of meanings only provides a limited set for ranking in a certain way, so no identical ranking should be expected. Based on Kendall's tau test, we interpret the fact that the mean τ is a positive value, so the

Uralic data rather confirms the stability of the empirically selected LJ meanings than deviates from it. Uralic does not stand out as a special outlier in comparison to the crosslinguistic sample, as a negative value of the mean τ would suggest.

4.2. Comparing the total basic vocabulary lists

The second research task was to answer the following question: Which of the basic vocabulary lists included in UraLex 2.0 is the most accurate, i.e. has the most stable meanings and the least borrowing-prone meanings?

We compared the unranked Swadesh100 and Swadesh207 lists, the ranked LJ, and the WOLD 401–500 lists with each other to examine their differences in borrowing-resistance for Uralic. We studied which list has the largest percentage of meanings from the cluster with the most stable meanings (cluster A) and from the most borrowing-prone meanings (cluster C). 45% of the Swadesh100 list was composed of highly borrowing-resistant meanings but 21% is borrowing-prone. The less-basic WOLD 401–500 has the lowest percentage of borrowing-resistant meanings (29%) and the highest percentage of more borrowing-prone meanings (33%) as can be expected. We hypothesized that the ranked Leipzig–Jakarta would be the most accurate list, but this hypothesis has to be rejected to a degree. 24% of the meanings in LJ meanings are borrowing-prone, a higher percentage than in Swadesh100, and 42% are stable (45% in Swadesh100). The rest of the meanings landed in the "mostly stable" cluster B. As a whole, the differences between the well-established lists are quite small. Furthermore, when one closely examines the meanings which were included in the LJ list but not in the Swadesh100 (see Section 4.3), the LJ has fewer "extremely" borrowing-prone meanings than Swadesh100 which, on the other hand, strengthens the reliability of LJ in capturing stable meanings. Logically, the longer the list, the more meanings of any kind find a place on the list, which means that Swadesh207 had the largest number of meanings from all clusters.

As the results of Swadesh's work and the WOLD survey suggest, the shorter, more carefully curated basic vocabulary lists tend to be more accurate for Uralic. The WOLD 401–500 falls outside of the core basic vocabulary. Our results are supported by a quantitative analysis on

tree-likeness using the TIGER algorithm and the UraLex 1.0 dataset: the core basic vocabulary lists reveal a stronger vertical tree-like signal of inheritance in comparison to the WOLD 401–500 (Syrjänen et al. 2021).

Our results show that even if there is some fluidity in the selection of the most basic meanings, the standardized lists fit Uralic relatively well. It may be more meaningful to identify and observe the behavior of individual problematic meanings than worrying about picking a "correct" total list.

4.3. Unexpected stability and borrowability of meanings

The third research task of this study was to find out which meanings behave differently in Uralic than assumed by the standardized lists.

We expected the WOLD 401–500 meanings to regularly appear among the most borrowable meanings. Still, 25 of the 99 meanings appeared surprisingly stable. For six meanings, the reflexes were words widely inherited across Uralic without semantic shifts, often being strict cognates. They are the best cases of truly stable meanings from WOLD 401–500. Sixteen meanings were rarely borrowed but, upon closer examination, the words for the languages were acquired through innovative word-acquisition strategies: derivation from different inherited stems (10 meanings), compounding, onomatopoeia, or analytic expressions (6). Even though both groups meet the criterion of not being borrowed, such meanings in Uralic do not really fit well with other criteria for basic vocabulary, namely morphemic simplicity and semantic neutrality. These violations together with the results of our clustering described in Section 4.2 align the Uralic data with the WOLD ranking. The WOLD 401–500 meanings seem to be more prone to multiple kinds of linguistic change, including borrowing.

The unexpectedly stable meanings also revealed a weakness in our method. Meaning slots with large amounts of missing data appear stable. This draws attention to the well-known general problem with (longer) basic vocabulary lists; some meanings are harder to define. In the future, clarifying the semantic components and language-specific connotations of the problematic meanings could help to correct for them at the data-collection phase.

Since our analysis only returned 25 surprisingly stable meanings, not much can be said about their semantic patterns. Three meanings

for (secondary) body parts ('armpit', 'heel', 'lung', 'molar tooth') and two body-related verbs ('to dream', 'to shit') from WOLD 401–500, confirm that body-related terminology is highly borrowing-resistant. Fourteen meanings are verbs, a category often among the less borrowable parts-of-speech (Matras 2009: 156–157).

It is possible that the surprisingly stable meanings are also affected by frequency. The criteria that frequent words should be used to denote basic meanings is implicitly accounted for the Swadesh lists and explicitly incorporated when establishing the WOLD ranks and the LJ list. It is assumed that stability increases together with high frequency (Haspelmath & Tadmor 2009: 15). However, the position of the WOLD 401–500 meanings in the WOLD data set implies that these meanings could be less frequent than higher ranks. It can be speculated that that the lower frequency of these meanings has made them less likely to appear in bilingual contexts facilitating borrowing. However, this is highly dependent on the unique contact situation between languages as low-frequency and specific words are borrowed as well (Backus 2014: 35). At the same time, studies on the continuum between code-switching and borrowing suggest that high frequency of new words makes it more likely for them to become a part of a monolingual person's vocabulary (e.g. Myers-Scotton 1993), while speakers' volition and their openness to accept new words play a role among other psycholinguistic factors (Backhus 2013: 36–37; Padraich & Roberts 2019). All these factors are necessarily integrated when basic vocabulary is borrowed as the precondition for intensive linguistic contact is bilingualism (Thomason & Kaufman 1988: 74–76). For our study as well as WOLD, a challenge is that comparable large frequency corpora are not available for the languages, which means that the exact effect of frequency in borrowing of basic vocabulary remains a subject for further study.

A smaller group of eight meanings appear more borrowing-prone than expected. Their reflexes were commonly borrowed in Uralic, whereas according to their occurrence in the Swadesh lists, especially Swadesh100, they should be borrowing-resistant. A pattern of independent borrowing events between donors and the languages in UraLex 2.0 can be identified. The borrowing events have resulted in a variety of loanwords, thus indicating that the eight meanings are truly unstable.

The loanwords were borrowed at different times in unique contact situations. Parallel borrowing from the same source language and within

Uralic can be found. This group reflects a continuum of well-known contacts with geographically neighboring languages. One meaning, namely ‘cold’, instead reflects a single borrowing event between Proto-Uralic (or possibly a younger Western Uralic entity) and an archaic Indo-European source (Koivulehto 2001, Grünthal 2025). This is exceptional as it is not common to have evidence of such deep-level contact. The ‘green-yellow’ adjectives are particularly interesting, because their reflexes include Indo-Iranian borrowings that were likely acquired into Western Uralic and then retained. Some languages have borrowings from their current neighbors representing older but also recent contacts. It may be that newer borrowings replaced older ones, but our data cannot confirm this or say how many turnovers might have occurred. The synonymy attribute, too, is interesting: if the meaning slot has both the old and newer borrowing, further research on the semantic nuances and contexts of use could give an insight into the turnover process in the meaning slots.

When the eight surprisingly unstable meanings are compared against the whole WOLD analysis (Haspelmath & Tadmor 2009), they belong to less borrowable meanings. In the context of Uralic and our study, they are highly borrowable. In other words, they are not very “basic” for Uralic specifically. The eight meanings also underline a difference between Swadesh100 and LJ in stability and the importance ascribed to the selection of meanings. When one inspects the total lists (see Section 4.2), the Swadesh100 appears as the most stable, because as it has more highly borrowing-resistant items than LJ. Still, the Swadesh100 also includes surprisingly borrowable meanings not found in LJ. The Leipzig–Jakarta list thus seems to be a stronger candidate after all for a well-rounded basic vocabulary list, because all the LJ meanings can be considered as “highly stable” or “mostly stable”. Swadesh’s intuitively selected meanings lack statistical support for stability, whereas the meaning selection for LJ is supported by a large-scale empirical study. Most meanings included in both Swadesh100 and LJ (conservative cluster A) behave as expected, i.e. they are the least likely to be borrowed. It can be claimed that both lists capture basic vocabulary well.

4.4. Challenges and conclusion

The WOLD language sample reflects, as all datasets unavoidably do, biases which can affect the meaning rankings in the background. WOLD is biased towards languages whose loanwords had already been studied in detail (Haspelmath & Tadmor 2009: 3, 55). Furthermore, the WOLD meanings can also be biased towards meanings relevant for European languages (Kruspe 2009: 671). The only Uralic language present in WOLD, Kildin Saami (Rießler 2009), is a "High Borrower" language (Haspelmath & Tadmor 2009: 56). Uralic basic vocabulary in general features an average or medium amount of borrowing in comparison to the large WOLD sample. A close examination of a six-language sample gives an average of 24% borrowing in basic vocabulary for Uralic (De Heer et al. 2024). In WOLD, average borrowers show 25–50% of borrowing and low borrowers 10–25%, and based on this, Uralic can fall between medium- and high-borrower languages (Tadmor, 2009: 57). In our case, the WOLD's bias toward high–medium borrower languages works to our advantage, because a basic vocabulary list based on such languages is likely realistic for Uralic as well.

We acknowledge that the amount of borrowing reflects the amount of etymological literature available. For example, more etymological research has been done on Finnic and Saamic than on Selkup. For the middle group, the contacts between Turkic and Uralic are well studied and there is an extensive tradition of research on loan etymologies. In the case of eastern Uralic languages, the study of loanwords is currently underdeveloped even though some important new results have been achieved in recent years. The UraLex 2.0 data and the handful of loanwords identified for eastern Uralic might represent a problem, because basic vocabulary lists may contain unidentified borrowings, including older ones. A language may also have adopted fewer loanwords during its history, or borrowings might have been replaced with newer loanwords from unidentified languages (Aikio 2002: 51). It is also possible that borrowings acquired in the distant past have not yet been recognized as such (Kaheinen 2023: 10). For instance, the hypothesis of early contact between pre-Proto-Tocharian and pre-Proto-Samoyedic has gained more support (see e.g. Warries 2022, 2025), which could lead to the discovery of more borrowings. Another possibility complicating the situation with Samoyedic that Khanina (2022: 85) has presented, is

that the relationships between Samoyedic language varieties show significant fluidity. This in turn might make it difficult to distinguish true established borrowings from more occasional code switches. At present, borrowing information for basic vocabulary in the eastern Uralic languages is lacking, but we accounted for this bias in our analysis (see Section 2.2).

Basic vocabulary lists are approximations of stable, borrowing-resistant meanings. In light of our results, these approximations capture a noticeable proportion of such items for Uralic too. After all, the differences in accuracy between the lists are not large. As the fit of the available basic vocabulary lists, especially LJ and the Swadesh100, is relatively good, we have chosen not to select and further curate a Uralic-specific basic vocabulary list. This is also to emphasize the cross-linguistic comparability of the UraLex 2.0 data. Furthermore, with future progress on loanword studies it is possible that loanword information for basic vocabulary could change even for popular lists. For the existing lists examined, it is important to understand the criteria according to which the basic meanings are selected, the overlap of the lists and the effect of borrowing in these lists. Claims such as universal availability of meanings or borrowing resistance should not be taken for granted even when well-known basic vocabulary lists are selected as research material.

Acknowledgments

We thank Timo Rantanen (University of Turku) for providing speaker-area information and Christopher Culver for reviewing the English language of the manuscript.

References

Aikio, Ante. 2002. New and old Samoyed etymologies. *Finnisch-Ugrische Forschungen* 57. 9–57.

Backus, Ad. 2014. A usage-based approach to borrowability. In Eline Zenner & Gitte Kristiansen (eds.), *New Perspectives on Lexical Borrowing: Onomasiological, Methodological and Phraseological Innovations*, 19–40. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9781614514305.19.

Bastin, Yvonne, André Coupez & Michael Mann. 1999. Continuity and divergence in the Bantu languages: Perspectives from a lexicostatistic study. *Journal of African Languages and Linguistics* 26(1). 111–115.

Campbell, Lyle. 1999. *Historical linguistics. An introduction.* Edinburgh: Edinburgh University Press.

Campbell, Lyle, & William Poser. 2008. How to show languages are related: the methods. In Lyle Campbell & William Poser (eds.), *Language classification: History and method*, 162–223. Cambridge: Cambridge University Press.

Chang, Will, Chundra Cathcart, David Hall & Andrew Garrett. 2015. Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language* 91(1). 194–244. https://doi.org/10.1353/lan.2015.0005.

Cronhamn, Sandra. 2018. *Quantifying Loanwords: A study of borrowability in the finnish lexicon.* General Linguistics Master's Programme: Language and Linguistics Lund University. MA-thesis.

Csúcs, Sándor. 2005. *Die Rekonstruktion der permischen Grundsprache*. (Bibliotheca Uralica 13). Budapest: Akadémiai Kiadó.

Dybo, Anna. 2013. *Etimologičeskij slovar' tyurkskikh jazykov. Tom 9 (dopolnitel'nyj): Etimologičeskij slovar' bazisnoj leksiki tyurkskikh jazykov*. (Tyurkskie slovari). Astana: TOO.

De Heer, Mervi, Mikko Heikkilä, Kaj Syrjänen, Jyri Lehtinen, Outi Vesakoski, Toni, Suutari, Michael Dunn, Urho Määttä & Unni-Päivä Leino. 2021. Uralic basic vocabulary with cognate and loanword Information. Zenodo. https://doi.org/10.5281/zenodo.4777568.

De Heer, Mervi, Rogier Blokland, Michael Dunn & Outi Vesakoski. 2024. Loanwords in basic vocabulary as an indicator of borrowing profiles. *Journal of Language Contact* 16 (1). 54–103. https://doi.org/10.1163/19552629-bja10057.

Dellert, Johannes, Thora Daneyko, Alla Münch, Alina Ladygina, Armin Buch, Natalie Clarius, Ilja Grigorjew, Mohamed Balabel, Hizniye Isabella Boga, Zalina Baysarova, Roland Mühlenbernd, Johannes Wahle & Gerhard Jäger. 2019. NorthEuraLex: A wide-coverage lexical database of northern Eurasia. *Language Resources and Evaluation* 54(1). 273–301. https://doi.org/10.1007/s10579-019-09480-6.

Dolgopolsky, Aharon. 1986. A probabilistic hypothesis concerning the oldest relationships among the language families of northern Eurasia. In Vitalij V. Shevoroshkin (ed.), *Typology, relationship and time. A collection of papers on language change and relationship by Soviet linguists*, 27–50. Ann Arbor: Karoma Publisher.

Embleton, Sheila. 1986. *Statistics in historical linguistics*. Quantitative linguistics (Volume 30). Bochum: Brockmeyer.

Erdal, Marcel. 2019. The Turkic-Mongolic lexical relationship in view of the Leipzig–Jakarta list. *International Journal of Eurasian Linguistics* 1(1). 78–97. https://doi.org/10.1163/25898833-12340005.

Field, Fredric. 2002. *Linguistic borrowing in bilingual contexts*. (Studies in Language Companion Series 62). Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.62.

Greenhill, Simon, Paul Heggarty & Russell Gray. 2020. Bayesian phylolinguistics. In Barbara Vance, Brian Joseph & Richard Janda (eds.), *The handbook of historical linguistics*, 226–53. Hoboken: John Wiley & Sons. https://doi.org/10.1002/9781118732168.ch11.

Fox, Anthony. 1995. *Linguistic reconstruction: An introduction to theory and method*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198700005.001.0001.

Grünthal, Riho. 2012. Baltic loanwords in Mordvin. In Riho Grünthal & Petri Kallio (eds.), *A Linguistic Map of Prehistoric Northern Europe.* (Mémoires de la Société Finno-Ougrienne 266), 297–343. Helsinki: Suomalais-Ugrilainen Seura.

Grünthal, Riho. 2025. Inherited features, grammatical change, and the influence of early language contacts in Finnic. In Björn Wiemer, Peter Arkadiev, Petar Kehayov & Rogier Blokland (eds.), *Convergence and Divergence in the Eastern Circum-Baltic Area. Volume 1: The Northern Part – A Synthetic View, especially on Finnic, and Case Studies* (Studia Typologica 33.1), 71–165. Berlin – Boston: De Gruyter Mouton. https://doi.org/10.1515/9783111165677-002.

Häkkinen, Jaakko. 2012. After the protolanguage: Invisible convergence, false divergence and boundary shift. *Finnisch-Ugrische Forschungen*, 61(1–3). 7–28. https://doi.org/10.33339/fuf.85674.

Haspelmath, Martin & Uri Tadmor. 2009. *Loanwords in the world's languages: A comparative handbook*. Leipzig: Walter de Gruyter. https://doi.org/10.1515/9783110218442.

Heggarty, Paul, Cormac Anderson, Matthew Scarborough, Benedict King, Remco Bouckaert, Lechosław Jocz, Martin Joachim Kümmel, Thomas Jügel, Britta Irslinger, Roland Pooth, Henrik Liljegren, Richard F. Strand, Geoffrey Haig, Martin Macák, Ronald I. Kim, Erik Anonby, Tijmen Pronk, Oleg Belyaev, Tonya Kim Dewey-Findell, Matthew Boutilier, Cassandra Freiberg, Robert Tegethoff, Matilde Serangeli, Nikos Liosis, Krzysztof Stroński, Kim Schulte, Ganesh Kumar Gupta, Wolfgang Haak, Johannes Krause, Quentin D. Atkinson, Simon Greenhill, Denise Kühnert & Russell Gray. 2023. Language trees with sampled ancestors support a hybrid model for the origin of Indo-European languages. *Science* (381). https://doi.org/10.1126/science.abg0818.

Hymes, Dell. 1983. Lexicostatistics and glottochronology in the 19th Century (with Notes Toward a General History). In Dell Hymes (ed.), *Essays in the History of Linguistic Anthropology*, 59–113. The Netherlands: John Benjamins Publishing Company.

Kaheinen, Kaisla. 2023. *Etymologia ex silentio: Nganasanin äännehistoria ja kielikontaktit*. Helsinki: University of Helsinki. Ph.D. thesis.

Khanina, Olesya. 2022. A history of Northern Samoyedic: adding details to the dialect continuum hypothesis. *Studia Uralo-altaica* 56. 77–94. https://doi.org/10.14232/sua.2022.56.77-94.

Holman, Eric, Søren Wichmann, Cecil Brown, Viveka Velupillai, André Müller & Dik Bakker. 2008. Explorations in automated language classification. *Folia Linguistica 42*(3–4). 331–354. https://doi.org/10.1515/FLIN.2008.331.

Koivulehto, Jorma. 2001. The earliest contacts between Indo-European and Uralic speakers in the light of lexical loans. In Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Consideration* (Suomalais-Ugrilaisen Seuran Toimituksia 42). 235–264. Helsinki: Suomalais-Ugrilainen Seura.

Kruspe, Nicole. 2009. 26. Loanwords in Ceq Wong, an Austroasiatic language of peninsular Malaysia. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the World's Languages: A Comparative Handbook*, 659–685. Leipzig: Walter de Gruyter. https://doi.org/10.1515/9783110218442.659.

Lehtinen, Jyri, Terhi Honkola, Kalle Korhonen, Kaj Syrjänen, Niklas Wahlberg & Outi Vesakoski. 2014. Behind family trees: Secondary connections in Uralic language networks. *Language Dynamics and Change* 4(2). 189–221. https://doi.org/10.1163/22105832-00402007.

List, Johann-Mattis. 2023. Inference of partial colexifications from multilingual wordlists. *Frontiers in Psychology* 14. https://doi.org/10.3389/fpsyg.2023.1156540.

Matisoff, James. 1978. *Variational semantics in Tibeto-Burman: The 'organic' approach to linguistic comparison*. Philadelphia: Institute for the Study of Human Issues.

Matras, Yaron. 2009. *Language contact*. Cambridge; New York: Cambridge University Press.

McMahon, April, Paul Heggarty, Robert McMahon & Natalia Slaska. 2005. Swadesh sublists and the benefits of borrowing: An Andean case study. *Transactions of the Philological* Society 103(2). 147–170. https://doi.org/10.1111/j.1467-968X.2005.00148.x.

McMahon, April & Robert McMahon. 2005. *Language classification by numbers*. Oxford: Oxford linguistics; New York: Oxford University Press. https://doi.org/10.1093/oso/9780199279012.001.0001.

Millar McColl, Robert Trask & Larry Robert. 2015. *Trask's historical linguistics.* 3rd Edition. New York: Routledge. https://doi.org/10.4324/9781315728056.

Myers-Scotton, Carol. 1993. *Duelling languages. Grammatical structure in codeswitching.* Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198240594.001.0001.

Nelson-Sathi, Shijulal, Johann-Mattis List, Hans Geisler, Heiner Fangera, Russell Gray, William Martin & Tal Dagan. 2010. Networks uncover hidden lexical borrowing in Indo-European language evolution. *Proceedings of the Royal Society of London B: Biological Sciences* (278). 1471–2954. https://doi.org/10.1098/rspb.2010.1917.

Padraic Monaghan, & Seán Roberts. 2019. Cognitive influences in language evolution: psycholinguistic predictors of loan word borrowing. *Cognition* 186. 147–58. https://doi.org/10.1016/j.cognition.2019.02.007.

Rankin, Robert. 2003. The comparative method. In Brian Janda & Richard Joseph (eds.), *The Handbook of historical linguistics*, 199–212. Hoboken: Wiley. https://doi.org/10.1002/9780470756393.ch1.

Rießler, Michael. 2009. Kildin Saami vocabulary. In Martin Haspelmath & Uri Tadmor (eds.), *The World Loanword Database*, 384–413. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://doi.org/10.1515/9783110218442.384.

Robbeets, Martine. 2020. Basic vocabulary in the Transeurasian languages. In Martine Robbeets & Alexander Savelyev (eds.), *The Oxford guide to the Transeurasian languages*, 645–659. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198804628.003.0037.

Rozhanskiy, Fedor & Mikhail Zhivlov. 2019. Votic and Ingrian core lexicon in the Finnic context: Swadesh lists of five related varieties. *Linguistica Uralica* 55(2). 81–108. https://doi.org/10.3176/lu.2019.2.01.

Sammallahti, Pekka. 1998. *The Saami languages: An introduction*. Karasjok: Davvi girji cop.

SSA = Itkonen, Erkki (ed.). 1992. *Suomen sanojen alkuperä: Etymologinen sanakirja. 1, A–K* (Suomalaisen Kirjallisuuden Seuran Toimituksia 56, Kotimaisten kielten tutkimuskeksuksen julkaisuja 62.). Helsinki: Suomalaisen Kirjallisuuden Seura, Kotimaisten kielten tutkimuskeskus.

Starostin, George. 2010. Preliminary lexicostatistics as a basis for language classification: A new approach. *Journal of Language Relationship* 3. 79–116.

Starostin 1998–2005 = Starostin, Sergei. 2016. *The global lexicostatistical database.* https://starling.rinet.ru/new100/main.htm.

Swadesh, Morris. 1952. Lexico-statistic dating of prehistoric ethnic contacts: With special reference to North American indians and eskimos. *Proceedings of the American Philosophical Society* 96(4). 452–463.

Swadesh, Morris. 1955. Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics* 21(2). 121–137. https://doi.org/10.1086/464321.

Syrjänen, Kaj, Terhi Honkola, Kalle Korhonen, Jyri Lehtinen, Outi Vesakoski & Niklas Wahlberg. 2013. Shedding more light on language classification using basic vocabularies and phylogenetic methods: A case study of Uralic. *Diachronica* 30(3). 323–352. https://doi.org/10.1075/dia.30.3.02syr.

Syrjänen, Kaj, Jyri Lehtinen, Outi Vesakoski, Mervi de Heer, Toni Suutari, Michael Dunn, Urho Määttä & Unni-Päivä Leino. 2018. *Lexibank/Uralex: UraLex basic vocabulary dataset*. Zenodo. https://doi.org/10.5281/zenodo.1459402.

Syrjänen, Kaj, Luke Maurits, Unni-Päivä Leino, Terhi Honkola, Jadranka Rota & Outi Vesakoski. 2021. Crouching TIGER, hidden structure: Exploring the nature of linguistic data using TIGER values. *Journal of Language Evolution* 6(2). 99–118. https://doi.org/10.1093/jole/lzab004.

Tadmor, Uri. 2009. III. Loanwords in the world's languages: Findings and results. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the World's Languages: A Comparative Handbook*, 55–75. Berlin, New York: De Gruyter, Mouton. https://doi.org/10.1515/9783110218442.55.

Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press. https://doi.org/10.1525/9780520912793.

Urban, Matthias. 2022. Red, black, and white hearts: "Heart", "Liver", and "Lungs" in typological and areal perspective. *Linguistic Typology* 26(2). 349–74. https://doi.org/10.1515/lingty-2021-2081.

Warries, Abel. 2022. Towards a new comparison of the pre-Proto-Tocharian and pre-Proto-Samoyed vowel systems. *Indo-European Linguistics* 10(1). 169–213. https://doi.org/10.1163/22125892-bja10022.

Warries, Abel. 2025. *Tocharian and Samoyed: on the question of Uralic substrate influence in Tocharian.* Leiden: Leiden University. Ph.D. thesis.

WOLD = Haspelmath, Martin & Uri Tadmor (eds.). 2009. *World loanword database*. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wold.clld.org/.

Woods, Anthony, Paul Fletcher & Arthur Hughes. 1986. *Statistics in language studies* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Kokkuvõte. Mervi de Heer, Luke Maurits: Uurali keelte baassõnavaraloendite tähenduste stabiilsus. Baassõnavaraloendite tähendustelt eeldatakse stabiilsust. Erialakirjanduses on pakutud välja mitmesuguseid erinevatel kriteeriumitel põhinevaid baassõnavaraloendeid, mille kõigi puhul on võtmekriteeriumiks stabiilsus. Baastähenduste stabiilsust ja seda, mil määral populaarsed nimekirjad neile sobivad, hinnatakse aga harva tervete keelerühmade puhul. Käesolevas artiklis defineerime stabiilsust kui vastupanu laenamisele ja uurime kvantitatiivselt, kuivõrd altid on tähendused populaarsetes baassõnavaraloendites uurali keelkonna puhul laenamisele. Täpsustame tähenduste järjestamise metoodikat, arvutades iga tähenduse laenamise tõenäosuse ja rühmitades need laenatavamate ja vähem laenatavate tähenduste rühmadesse. Meie kvantitatiivne analüüs annab baassõnavara piires parema lahutusvõime. Kuigi baassõnavaraloendid sobivad üldiselt hästi uurali keelkonda, näitame, et mitte kõik põhitähendused ei ole võrdselt stabiilsed ega tingimata baastasemel selle sõna kõige kitsamas tähenduses.

Märksõnad: baassõnavara, Swadeshi nimekiri, tähendusloendid, laensõnad, laenamine, kvantitatiivsed meetodid