Etymological research has a long and fruitful history in Finland, and Finnish may in fact be the most thoroughly etymologically studied language outside the Indo-European language family. Several etymological dictionaries of Finnish have been published, most notably the multi-volume *Suomen kielen etymologinen sanakirja* (‘the etymological dictionary of Finnish’) (SKES; 1955–1981) and its successor *Suomen sanojen alkuperä* (‘the origin of Finnish words’) (SSA; 1992–2000). These works are primarily directed at the scholarly community, and until recently Kaisa Häkkinen’s brief *Etymologinen sanakirja* (1987) has been the only reference written with the non-expert reader in mind. The work under review, *Nykysuomen etymologinen sanakirja* (‘the etymological dictionary of modern Finnish’, henceforth NES) by the same author, is an updated and greatly expanded version of this earlier work, which makes it the third comprehensive etymological dictionary of Finnish.

NES is a single hardcover volume of 1633 pages. Due to the size of the dictionary it would have been a good idea to publish it in two volumes, which would have made it both less clumsy to use and more resistant to wear. One expects a reference work of this kind to last in heavy use, but the binding of NES regrettably does not promise the book a long life; the back of the review copy got torn through wear already after a year of use.

The word articles cover 1532 pages, in addition to which there is a brief introduction (p. 6–16), explanations of the most common special characters (p. 17), references (p. 1551–1570), an index of inventors of neologisms (p. 1571–1574), and an index of words discussed in other word articles but lacking one of their own (p. 1575–1633). On the inside of both the front and the back covers there is a map showing the distribution of the Finno-Ugric (Uralic) languages, a useful inclusion for readers unacquainted with the language family. Certain details of the map could have been improved, though: the speaking area of Inari Saami is missing, and the Saami and Samoyed languages have been indicated with single symbols despite of their deep divergence.

On the first skim NES leaves a very favorable impression. It is pleasing to note that Häkkinen has disengaged herself from the Finno-Ugrianist jargonism characteristic of the earlier etymological dictionaries.

Reviewed by Ante Aikio
of Finnish. The word articles are written in a clear and informative style, without long word lists, sentence fragments or heavy use of abbreviations. This is a standard practice in good etymological dictionaries elsewhere in the world (e.g. Kluge 1995; Magnússon 1989), and it is praiseworthy that Häkkinen has brought the tradition to Finland.

1. The selection and citing of material

As the title of NES tells, the lexical material chosen in the dictionary represents the vocabulary of ‘modern Finnish’. The number of word articles in the dictionary is somewhat over 6000, which Häkkinen says to accord with Erkki Itkonen’s estimate of the number of underived word-stems in Finnish – a figure which does not include young loanwords and so-called ‘learned words’. Even so, the 6000 word articles include also new loans and internationalisms as well as derivatives and compounds, whereas the bulk of dialectal and obsolete vocabulary has been left out. The selection of material thus differs substantially from the previous etymological dictionaries.

The inclusion of recent loans and internationalisms is an excellent choice, as they, too, form an integral part of the ‘modern Finnish’ lexicon. However, this good idea seems to have been carried out in a somewhat arbitrary way: for instance, abstrakti ‘abstract’, adressi ‘address, petition’ and aforsmi ‘aphorism’ are included, whereas absoluuttinen ‘absolute’, adoptio ‘adoption’, aerobic ‘aerobics’, agentti ‘agent’ and aggressio ‘aggression’ are not. It is not easy to see what kind of criteria have been applied here, and in any case a substantial part of this type of vocabulary is missing. Moreover, some entirely common words have slipped out as well, such as aavikko ‘desert’, ahtaa ‘to cram, stuff’ (this missing word article is even referred to on p. 27), katkarapu ‘shrimp’, kiihtyä ‘to accelerate; to get upset’, no ‘well (discourse particle)’, nääntyä ‘to starve’, sietää ‘to endure’, siika ‘whitefish’, siili ‘hedgehog’, siima ‘line (for fishing)’, siipi ‘wing’, vala ‘oath’ and väsyä ‘to get tired’. Also helmikuu ‘February’ and huhtikuu ‘April’ are missing, even though tammikuu ‘January’ and maaliskuu ‘March’ are included. Such shortcomings suggest that the book was perhaps too hastily edited.

Some remarks can also be made on the citing of linguistic material. Häkkinen has made a few amendments to the standard Finno-Ugric phonological transcription, including the introduction of the grapheme <õ>, which stands both for a reduced vowel /œ/ or /œ:/ (e.g. in Mari and Khanty),
and a mid central unrounded vowel /e/ (e.g. in Komi and Udmurt), as well as the replacement of /i/ (a high central unrounded vowel) with the letter <y>. While the idea of getting rid of unnecessary diacritics is in itself sensible, the justification given for these changes is somewhat curious: Häkkinen states that she has “used as simple an orthography as possible, which guides also the reader unfamiliar with special characters to read the word approximately correctly” (p. 12).¹ One can wonder whether this is a relevant or even a possible goal; and in any case ó is just another special character to the non-expert, who will no doubt also find it hard to associate y with a central unrounded vowel, as the same letter stands for a front rounded vowel in the Finnish orthography.

Even so, Häkkinen’s changes to the standard transcription are small and cause no difficulty for specialists. It is more regrettable that there are occasional errors in the cited forms, a good example of which is provided by the Saami data. NES applies the standard spelling of North Saami instead of the complicated Nielsen’s orthography common in earlier etymological references, which would be a good choice, were it not that the forms include numerous transliteration mistakes. Examples include guoddu pro guottu ‘tree stump’ (p. 351), gaehpid pro geahppat ‘light (not heavy)” (p. 400), leaggji pro leadji ‘heap’ (p. 652), njuoskas pro njuoskkas ‘wet’ (p. 768), čavggi pro čávggi ‘beam in a sod hut’ (p. 1132), sagnat pro sakya ‘to thaw’ (p. 1223) and dárpmi pro dármbi ‘energy, vigor’ (p. 1279); this list is far from exhaustive. Such mistakes are not confined to Saami — for example, Komi tšeltny and Udmurt tšeltnyny ‘to throw, etc.’ (p. 1210) should read tšöltmy and tšöltyn (= cěltmi and cěltiŋi in standard transcription), Hungarian hájt ‘to throw’ (p. 323) should read hájit, and the word turo ‘ski’ (p. 1203) is not Nenets but Enets instead.

2. The principles of etymology

The introduction of NES provides a lot of information useful to non-expert readers: Häkkinen briefly discusses the principles and criteria of etymological research, the development of the Finnish literary language, as well as the stratification of the Finnish lexicon and the most important sources of loanwords. Of particular interest here are the theoretical and methodological issues touched upon, as they also have a bearing on the content of the word articles. In Häkkinen’s view etymology operates with

¹ This quote, as well as the ones that follow, are my translations from Finnish.
three basic criteria: 1) the distribution, 2) the structure and phonological shape, and 3) the meaning of the word. As for the hierarchy of these criteria, she writes that “especially when evaluating the true age of the oldest words, the central criterion is the distribution. As a rule of thumb one may regard that the wider the word is attested in the cognate languages of Finnish, the older the word in question is.” (p. 7; cf. 1990: 24–67)

Two remarks are in order here. First, distribution as such actually reveals little of the age of a word, because loanwords can also show extensive distributions. Instead, it is phonological criteria that are used to establish the age of words: old inherited words show regular sound correspondence between related languages whereas loanwords display different correspondences. Second, ‘distribution’ cannot really count as an etymological criterion because it is a research result instead: one cannot know the distribution of any word before it has been established through etymological study. 

Admittedly, this is partly acknowledged by Häkkinen as well, as she writes that “the criteria concerning phonological and morphological structure are needed when it is determined which words belong together etymologically...” (p. 7) and that “[the distribution] cannot be considered decisive if it seems to be at odds with the phonological criteria” (p. 8). Even so, the discussion on the application of phonological criteria is a bit confusing: Häkkinen states that one cannot reliably determine the age of loanwords on the basis of their phonological shape alone as unassimilated loans such as farao ‘pharaoh’, graniitti ‘granite’ and strutsi ‘ostrich’ already occurred in the earliest Finnish literary language (p. 9). Because etymology usually operates on a timescale of millennia rather than centuries, farao and the like count as young words regardless of whether they were present in “old” literary Finnish or not. What is more, the reader is also told that such words as kinkku ‘ham’, penkki ‘bench’, sielu ‘soul’ and ralli ‘rally’ may look like “autochthonous” words and could not be deduced as borrowings on the basis of their phonological shape alone (p. 15), even though it is textbook knowledge of Finnish historical phonology that the earlier language stages had no unstressed rounded vowels, geminate sonorants or three-consontant clusters (see e.g. Häkkinen 2002: 64–69).
3. Loanwords

As the bulk of Finnish vocabulary consists of loanwords, it is natural to deal with them first. Recent loanwords and internationalisms are treated well in NES, and a few word articles also provide welcome background information. For instance, the reader looking up bussi ‘bus’ is certainly interested in knowing that buses were introduced as a means of public transport in Helsinki in the 1920s, and prior to this bussi was used as a slang word for streetcars, which were originally horse-drawn.

Occasionally, though, one would have hoped for a more exact wording. For instance, of iglu ‘igloo’ NES states that “originally the word derives from the Inuit or Eskimo language word igdlu.” Eskimo is not a single language, and Inuit is a subbranch of the Eskimo group, which also includes the Yupik languages spoken in Alaska and on the Chukchi Peninsula. In the various individual Eskimo languages this word appears in quite different forms such as iylu, illu, ittiq, øyluq, lu, etc. The exact source of iglu seems to be the form iylu attested both in North Greenlandic and in certain Inuit languages of Alaska and Canada (Fortescue, Jacobson & Kaplan 1994: 112).

Older loanwords from Germanic, Baltic, Aryan etc. are also generally treated well in NES. Often a reconstruction of the loan original is provided, as well as a few examples of its reflexes in the attested languages. However, in some cases the treatment could have been improved by presenting more information on the phonological development of the Finnic forms as well. Finno-Ugric reconstructions are routinely given in NES for old inherited words, but not in the case of old loans from Baltic and Germanic. For example, the connection between Finnish halla ‘frost’ and Lithuanian šalnà ‘frost; rime’ would no doubt seem more evident to the non-expert reader if he was told that the Finnish form has developed from Pre-Finnic *šalna via the regular changes *š > h and *ln > ll.

In general loanwords receive a sound treatment in NES. The earlier dictionary SSA tends to be somewhat overcritical toward loan explanations in general, and it is pleasing to note that several convincing loan etymologies that SSA classifies as uncertain are accepted in NES; these include, for instance, the Germanic etymologies of Finn. nahka ‘skin, hide’, paasi ‘stone slab’, rasia ‘box’, runko ‘trunk’, tuppi ‘sheath’, uskoa ‘to believe’, vihjata ‘to hint’ and vihko ‘bundle, sheaf; notebook’. In a couple of cases, though, perhaps too cautious an attitude has been adopted in NES as well. For instance, there is hardly any reason to doubt that Finnish vyyhti
‘skein’ (~ Votic *vehsi id.) < *vepti is a borrowing from Proto-Germanic *wefti- ‘weft’ (> English *weft, Swedish väft). This etymology is considered “probable” in SSA, whereas according NES the word only “possibly” derives from Germanic. Besides, the attested English form is misspelled as *veft in NES.

4. Finno-Ugric inheritance

While the bulk of Finnish vocabulary consists of loans and other innovations, a few hundred of its core lexical items are inherited from earlier Finno-Ugric proto-language stages. Such old vocabulary has for long been a central object of etymological study, and the general picture of shared lexical material in the Finno-Ugric family seems quite clear by now, even though new etymological equations are still occasionally discovered. Even so, more important advances in this subfield have recently been reached through the critical reassessment of the corpora of old etymologies rather than the discovery of new ones. During the past few decades the understanding of Finno-Ugric historical phonology has become more precise especially due to the studies by Janhunen (1981) and Sammallahti (1988). At the same time the number of Finno-Ugric etymons that are considered reliably reconstructed has decreased rather than increased, as many comparisons that were considered well-founded in earlier research have turned out to involve irregularities after all.

How are these developments reflected in NES? Häkkinen’s position can be characterized as rather traditionalistic: even though she states that “the earlier etymologies have been critically gone through and... amended when needed” (p. 6), in reality NES often just repeats what is said in SSA or *Uralisches etymologisches Wörterbuch* (UEW; 1988). Also the majority of proto-language reconstructions given for Finno-Ugric words are taken directly from UEW, which is a bad choice because there has been notable progress in phonological reconstruction since the publication of this dictionary. For instance, Janhunen (1981) and Sammallahti (1988) have convincingly argued that a consonant *x* as well as a high back unrounded vowel *i* must be reconstructed for Proto-Uralic, both of which are missing from UEW’s reconstructions.

In fact, UEW also ignores much of the etymological research conducted after the 1960s, and because of this the dictionary was regrettably outdated already at the time of its publication in 1988. This makes it a poorly suited primary source for Finno-Ugric etymologies, and
now many outdated comparisons continue their life in NES. Let us give a couple of examples:

- Finn. savi ‘clay’ is equated with Erzya Mordvin šovoň, Mari šun, Komi šun, Mansi šul’ and Khanty sówî (all with the same meaning ‘clay’), and the Finno-Ugric proto-form is reconstructed as *šawe. The sound correspondences between these words are completely irregular, and the presented reconstruction has no basis. Only the Mordvin, Mari and Komi forms are doubtlessly etymologically linked, and even in their case borrowing in one direction or the other is possible.

- Finn. vaalea ‘light-colored’ is equated with Udmurt val’ ‘shine, shimmer; shining’, Komi vol’alny ‘to shine, glitter’ and Mansi wol’gi ‘to shine’, and the Finno-Ugric proto-form of these words is reconstructed as *wał’V. The vowel correspondences are entirely irregular, and the same applies to the consonant correspondence l ~ l’. According to the present view there are no reliable etymologies supporting the reconstruction of a phoneme *l’ in Finno-Ugric (Sammallahti 1988: 491).

Comparisons which are demonstrably wrong are especially common in the case of Samoyed languages. The following serve as examples:

- Tundra Nenets ngawey° ‘brain’ is presented as an uncertain cognate of Finn. aivot ‘brain’. The word is actually a derivative of ngaewa ‘head’, which is cognate with Finnish oiva ‘good, splendid’ (cf. s.v. aivot, oiva).

- Nganasan mou ‘earth’ is presented as the cognate of Finn. maa ‘earth, land, ground’. However, it derives from Proto-Samoyed *måjà ‘earth, land’ (Janhunen 1977: 85), which is actually the cognate of Finn. muta ‘mud’ < *muð’a (Aikio 2002: 22–23).

- Selkup mäkte and Kamas mekte ‘tussock’ are given as cognates of Finn. mätäš id., and these are claimed to derive from Proto-Uralic *mäkte. This equation is phonologically unacceptable, because Proto-Uralic *k has regularly disappeared in Proto-Samoyed adjacent to obstruents (*t, *c, *s, *š): one would expect *mäkte to have developed into Selkup *mäte etc. (Janhunen 1981: 251).

One could have avoided repeating such mistakes of earlier etymological dictionaries by making more extensive use of newer primary studies. References to Janhunen (1981) and Sammallahti (1988), both fundamental works in Finno-Ugric historical phonology, are surprisingly rare in NES. In fact, as Janhunen and Sammallahti have already to a significant extent done the task of critically going through the earlier Finno-Ugric etymologies and
amending them as needed, it is not easy to see why Häkkinen has chosen to ignore these results and to start over with the same task.

Even though several outdated Finno-Ugric etymologies are included, references to many newer word comparisons are missing. NES does not mention, for instance, that according to Janhunen (1981) Finn. mukaan, mukana ‘along with’ are fossilized case forms of the Proto-Uralic noun *muka ‘back’, Finn. anoppi ‘mother-in-law’ is originally a compound consisting of Proto-Uralic *ïna ‘mother-in-law’ and *ippi ‘father-in-law’, and Finn. kantaa ‘to carry’ is a causative derivative of Proto-Uralic *kani- ‘to go away’. More recent etymological equations unmentioned in NES include those presented for Finn. muta ‘mud’ (Abondolo 1996: 28–29), lähi- ‘near’ (Helimski 1999), sivu ‘side’ (Kulonen 2001), juoda ‘to drink’, kaiva ‘to dig’, kääriä ‘to wrap’, kätkeä ‘to hide’, lämmin ‘warm’, nisä ‘teat’, sota ‘war’, salama ‘lightning’, seistä ‘to stand’ and täysi ‘full’ (Aikio 2002).

5. Sound-symbolic vocabulary

The origin of sound-symbolic vocabulary (a.k.a. ‘descriptive words’ in Finland) has always been one of the hardest nuts to crack in etymology. Irregular phonological variation, semantic instability and frequent associations and contaminations with other words make studies in this lexical domain notoriously difficult. Hence, it is not surprising that etymologists have often been content with just labeling such words as “descriptive” without analyzing them in detail. Recently, however, sound-symbolic vocabulary has received increasing attention in Finnish etymological studies.

Compared to the recently published SSA, where numerous words are characterized as ‘descriptive’ or ‘onomatopoetic’ without further argumentation, NES has taken a step forward. Since such ‘descriptive’ etymologies are often based on loose criteria and subjective impressions (see e.g. Koivulehto 2001; Aikio 2001; Nikkilä 2002; Mikone 2002: 33), it is excellent that Häkkinen has adopted a somewhat more cautious attitude: a number of explanations of this kind are treated with suspicion or even left unmentioned in NES, including those SSA proposes for kaapata ‘to seize, snatch’, kumpu ‘hill, hillock’, nuttu ‘jacket’, riippua ‘to hang’, tuoksua ‘to smell good’, tuppi ‘sheath’ and vihko ‘bundle, sheaf; notebook’.

Even so, there would still be room for improvement, as NES is not completely free from dubious descriptive etymologies either. For instance,
it is not easy to see why sarastaa ‘to shimmer, dawn’ should be “seen most likely as an autochthonous descriptive word”, how päre ‘shingle, chip’ could be “a derivative of a word-stem imitating a buzzing or a creaking noise”, or on what grounds hiipää ‘to sneak’ is interpreted as “a descriptive formation that describes the dragging noise caused by sneaking” (isn’t the purpose of sneaking to avoid making any noise?). Even a couple of new suggestions of this type are put forward, a surprising example of which is the explanation proposed for Finn. murhe ‘sorrow’ and murehtia ‘to grieve’: “According to a simple but entirely possible explanation the verb murehtia would be a derivative of the same sound-imitative root as murista [‘to growl’]... A person can express sorrow with lamentations that resemble growling.”

As with the assessment of proposed Finno-Ugric etymologies, one gets the impression that something is left half done. It is obviously an improvement that a number of problematic descriptive suggestions have been left out, but it is inconsistent at the same time to advocate other equally implausible explanations. It would seem justified to categorically reject all descriptive etymologies as long as they are not based on any explicitly defined and justified criteria. Loose guesswork will not advance the research, and more sophisticated methods should instead be developed for the etymological analysis of sound-symbolic and expressive vocabulary. Only recently there have been some real advances in this field; one can mention Nikkilä’s (1998; 1999) studies on ‘suffix replacement’ and ‘sporadic consonant alterations’ in Finnish, as well as Jarva’s (2003) monograph which sheds much light on expressivization of loanwords.

6. Words of unknown origin

In every language there remains a body of words of whose origins can never be explained, regardless of how thoroughly etymological studies are conducted. In light of this, it is a bit curious that the previous etymological dictionaries of Finnish have been reluctant to admit that the origin of any Finnish word is unknown. To avoid this, two strategies have usually been employed: either the lexical data is listed without stating anything of the etymological background, or “descriptive” origin is proposed (see above). In contrast to this, it is an improvement that NES often explicitly states that the word is of unknown etymology, as in the case of anastaa ‘to deprive, rob’, hullu ‘crazy, mad’, impi ‘virgin (poetic)’, iskeää ‘to strike’, loppu ‘end’, mekko ‘dress’, neuvo ‘advice’, nuotio ‘campfire’, outo ‘strange’,
pettää ‘to deceive, betray’, as well as in many others. This principle has not been applied entirely consistently, though: there are many times more words of unknown origin in Finnish than are classified as such in NES.

What, then, is said about the origin of the rest of this vocabulary? This time a reduction in the number of words of unknown origin has been achieved by applying the concept of ‘autochthonous words’ (“omaperäiset sanat”): words lacking an etymology are frequently classified into an ‘autochthonous’ lexical stratum. To give an example, Finnish kynä ‘pen; feather’ has cognates throughout the Finnic languages, but the further origin of the word is obscure. Hence, Häkkinen writes that “due to the lack of other explanations the word kynä is in any case considered an old autochthonous word.” There are innumerable similar word articles in NES.2

The problem with these numerous “autochthonous” words is that the concept itself is quite obscure. In NES one cannot find a clue how the term is supposed to be understood, but the following passage in Häkkinen’s earlier handbook on Finnish etymology sheds some light on the issue:

“One can maintain the working hypothesis that all words are ancient and autochthonous, until proved to the contrary. Applying all the basic criteria of etymological research, i.e. phonological form, meaning and distribution, one must then weigh whether there is reason to start suspecting another explanation.” (Häkkinen 1990: 88)

First, one must note that this “working hypothesis” is applied inconsistently in NES: some words are said to be of unknown origin whereas others are classified as old and autochthonous exactly because their origin is unknown. For instance, cognates of Finn. liha ‘meat’ are found only in Finnic languages, and hence according to NES the narrow distribution suggests that it is a loanword, even though no loan original has been

---

discovered. Because the distribution of *kynä* ‘pen; feather’ is equally narrow, one can only wonder why it is not suspected as a loanword as well.

Second, it is paradoxical to assume that the lack of an etymology would form an etymological explanation in itself. If there is no information about the origin of a given word, then its origin simply remains unknown; the term ‘autochthonous word’ seems to be a mere circumlocution for the lack of etymology. The easiest way out of this confusion would be to abandon the misleading term altogether; the constant references to ‘autochthonous’ origin are only likely to blur the readers’ understanding of how words can originate.

7. **In conclusion**

Even though critical remarks have been made above, this should not be understood to imply that NES is a failed work. On the contrary, the great majority of the word articles in NES are free of problems and reflect sound scholarship, and the dictionary presents a wealth of information on Finnish etymology in a clear and accessible form. This is already a substantial achievement in the popularization of Finno-Ugric comparative linguistics, and it is certainly appropriate that this work was nominated as a candidate for the respected Finnish *Tieto-Finlandia* prize, which is awarded annually for the best non-fiction book of the year.

On the other hand, an etymological dictionary cannot be reviewed as a purely popular work even if it has been authored primarily with the non-expert reader in mind. True, an etymological dictionary does not necessarily need to present many new research results, but it is still an extensive scholarly contribution to examine the entire body of relevant publications in order to judge which of the proposed etymologies can be considered plausible. One must note that the problems of NES – such as the misleading references to ‘autochthonous words’ or the less convincing ‘descriptive’ etymologies – seem to stem more from the traditions of Finnish etymology in general than from the author’s own ideas in particular. Hence one would have hoped that a somewhat more critical attitude had been adopted toward the explanations presented in earlier references.

Regardless of its flaws, one can readily recommend NES as a good reference for both linguists and non-experts interested in the origins of Finnish words. Even Finno-Ugric specialists may find it handy, especially due to the references given under each word article. As this work will
certainly find many readers over the next decades, one must hope that the next edition will be published with a more hard-wearing binding.

References


Contact information:

Ante Aikio
Giellagas Institute for Saami Studies
P.O. Box 1000
90014 University of Oulu
Finland
e-mail: ante.aikio@oulu.fi