



Name Variant Tools

Diane L. Richard wonders how many ways there are to spell a name

I CHALLENGE ANYONE TO FIND AN ANCESTOR whose surname is spelled exactly the same in every document. With so many clerks writing so many of our ancestors' documents for so long, we typically find our ancestors name spelled so many different ways!

I've seen documents where the clerk wrote "a" name at least five different ways and then the party signed it in yet a different way! Between lack of familiarity with a name, the use of phonetic spelling, and many other factors, there was no single way to spell anything (just look at 18th & 19th century court records, land records, etc), never mind a name!

Additionally, it was common for the clerks, and sometimes the individuals themselves, to abbreviate their forenames, e.g. Jas for James, Alexr for Alexander, etc. How does one most effectively search for all of the variants which might exist for a given name?

NameX

To better provide research-relevant sites, several genealogy databases, along with other users, are taking advantage of the NameX technology developed by Image Partners, www.imagepartners.co.uk/Welcome.aspx. The resulting forename and surname thesauruses, www.namethe-saurus.com/Thesaurus/Search.aspx, currently contain:

385 million variants identified for 5,929,000 Surnames

26 million variants identified for 1,246,000 Forenames

For example, I have been searching on some records for the surname Haight. NameX identified 105 different variants for that surname. Fortunately, a "match score" is provided which weights the various identified variants, and helps identify the more likely ones. Fortunately, in my research so far, I have only come across about 10 different spellings, and looking at this list really reminded me that I pretty much have to focus on names that have H*** and not count on much more! Though, I have frequently come across Hate as a possible spelling, and yet that ranked down the list lower than many other variants. And, we all know from our genealogy research that spelling variants are also highly dependent on region — the typical way to spell a name in one state, may differ from the typical way it is spelled elsewhere. Also, if it's not unusual for different branches of our own family tree to spell a surname differently, we really can't expect the clerks, census enumerators, etc. to do any better!

And, if the aforementioned tool doesn't seem to find the names you

seek, check out Kimberley Powell's About.com Genealogy column, **Top 10 Tips for Finding Alternate Surname Spellings & Variations**, http://genealogy.about.com/od/name_changes/tp/spellings.htm.

Though we've been talking about surnames, we do have some of the same issues with forenames — especially those that are often abbreviated. For example, the forename Squire, often based on Esquire, can be found in mid- to late-19th century NC records. Unfortunately, the NameX Forename Thesaurus doesn't even recognize Esquire as a forename. And, to be fair, this tool was developed by a UK-based company, so it does better with UK-based names. When I entered my own forename, Diane, I was amazed to see all the variants! I thought people pretty much

Original Name	Number of Variants	Match Score	Original Name	Number of Variants	Match Score
Haight	105	100%	Haight	105	100%
Haght	10	95%	Haght	10	95%
Haght	10	90%	Haght	10	90%
Haght	10	85%	Haght	10	85%
Haght	10	80%	Haght	10	80%
Haght	10	75%	Haght	10	75%
Haght	10	70%	Haght	10	70%
Haght	10	65%	Haght	10	65%
Haght	10	60%	Haght	10	60%
Haght	10	55%	Haght	10	55%
Haght	10	50%	Haght	10	50%
Haght	10	45%	Haght	10	45%
Haght	10	40%	Haght	10	40%
Haght	10	35%	Haght	10	35%
Haght	10	30%	Haght	10	30%
Haght	10	25%	Haght	10	25%
Haght	10	20%	Haght	10	20%
Haght	10	15%	Haght	10	15%
Haght	10	10%	Haght	10	10%
Haght	10	5%	Haght	10	5%
Haght	10	0%	Haght	10	0%

Example of NameX Surname Thesaurus result for Haight.

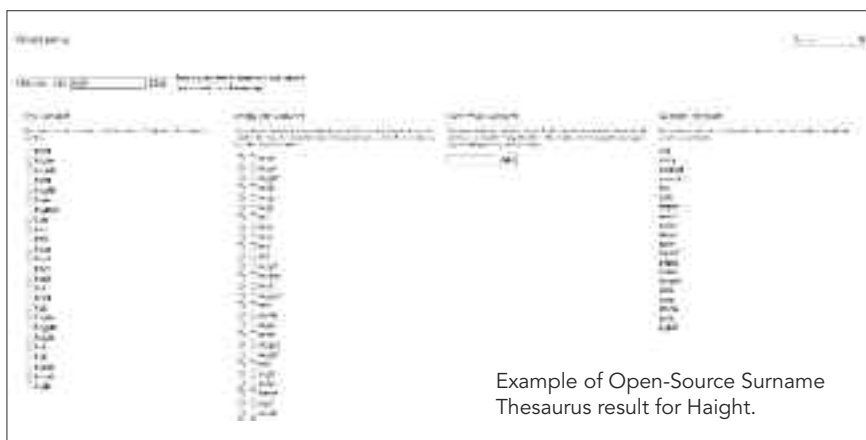
always spelled it Diane or Dianne; where did those other 87 options come from?

Open-Source Database of Name Variants

Along the same lines, Ancestry.com, BehindTheName.com, and WeRelate.org have created an open-source database of name variants that is free for any website or genealogy software developer to use. Learn more about the project at www.werelate.org/wiki/WeRelate:Variant_names_project.

This database was started by Ancestry.com and WeRelate working together to create an advanced algorithm for determining the level of similarity between two names. It was used to find similarly-spelled names for the 200,000 most-frequent surnames and 70,000 most-frequent given names in Ancestry's database. This includes every name that appears more than once in every 5 million names in Ancestry's database. On average, 26 variants were found for each surname, and 32 variants were found for each given name. In addition, BehindTheName.com donated their excellent list of given name variants. Other variants from the WeRelate community, The New American Dictionary of Baby Names, and A Dictionary of Surnames were also included.

Since this is an open-source project, the genealogy community is asked to help by adding variant spellings to the database. To see what variants are currently included, go to this page, www.werelate.org/wiki/Special:Names, and on the drop-down menu, select Surname or Given Name and enter the name of interest. For example, I revisited the Haight



Example of Open-Source Surname Thesaurus result for Haight.

surname discussed before, and you can see how the results are presented in a different format. Again, I was impressed with how many ways my forename, Diane, can be spelled!

Conclusion

Recognize that any tool, and not just the two mentioned, can't fully help us identify variants as found in indexes, abstracts or transcriptions where a person confused a P with a T or an F, or an m with n. Since these aren't name variants based on phonetics, they add another layer of complexity to identifying all the name variants one might consider when doing research.

And, as we know from George and Ira Gershwin's song, *Let's Call The Whole Thing Off...*

You say eether and I say eyether,
You say neether and I say nyther,
Eether, eyether, neether, nyther,
Let's call the whole thing off!
You like potayto and I like potahto,
You like tomayto and I like tomahto,
Potayto, potahto, tomayto, tomahto!

So, now let's imagine that poor clerk trying to figure out what you are saying, and then let's throw a heavy accent on it, a lack of uniform spelling rules, inconsistent penmanship, rushed writing (just think of all the notes you've

written to yourself when in a rush that you couldn't decipher later!), hard-to-use writing materials and sometimes it's amazing we can read anything!

Even with all the caveats I've presented about what these two tools can't do, and which we shouldn't expect them to do, they are really great tools to help us consider other ways that names might have been spelled. I've had too many discussions with clients who think that there is a right way to spell their name, until I show them the evidence of how many different ways I found their surname spelled.

These tools give me a head start in thinking beyond the modern conventional spelling as to how any name might have been spelled. I guess this means I won't be talking out loud to myself as much anymore, as I try to "sound out" how a name might have been said and then spelled! ©

DIANE L. RICHARD has been doing genealogy research for over 25 years. She is currently a professional genealogist and can be found online at www.mosaicrpm.com.

(Or scan the QR code with your mobile phone or table device.)

