The Predictability of Name Truncation: Factoring in Language Change*

Birgit Alber
Free University of Bozen-Bolzano
birgit.alber@unibz.it
https://orcid.org/0000-0002-6473-6098

Sabine Arndt-Lappe
Trier University
arndtlappe@uni-trier.de
https://orcid.org/0000-0001-7723-7151

Joachim Kokkelmans
Free University of Bozen-Bolzano
jkokkelmans@unibz.it
https://orcid.org/0000-0001-5568-2588

Received: May 3, 2024
Accepted: November 7, 2024
Published: February 27, 2025

Abstract

Truncation of proper names (e.g. English *Robert* > *Rob)* has been claimed to be a marginal or extragrammatical word-formation process due to the presumed unpredictability and irregularity of its outputs (Dressler & Merlini-Barbaresi 1994 et seq.). In this study, the hypothesis is pursued that the coexistence of multiple truncation patterns in a single language is in some cases due to language change: older patterns coexist with more recent ones, thus creating the false impression of random variability. This hypothesis is tested in an experiment where participants are asked to choose the most appropriate short name for a person of a certain age group in a dialogue. The experiment was conducted over the crowdsourcing platform AlpiLinK (<https://alpilink.it/>) for speakers of Romance and Germanic varieties of Northern Italy. The results of the experiment show that the truncation patterns tested for Italian and German names differ significantly in the age that participants associate with them. A second finding of this study is that diachronically close truncation patterns differ minimally in the grammars that correspond to each pattern in a formal model of truncation implemented as a typological system in Optimality Theory (Alber

* For useful comments the authors thank reviewers and editors as well as the audience at the Conference on Nominal Inflection and Word-Formation at the Phonology-Morphology Interface held at Barcelona, 25-26 January 2024. We furthermore thank Sabrina Raus for sharing test-items from her MA-thesis with us (Raus 2024). This research has been carried out within the PNRR research activities of the consortium iNEST (Interconnected North-East Innovation Ecosystem) funded by the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) Missione 4 Componente 2, Investimento 1.5 D.D. 1058 23/06/2022, ECS_00000043). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

& Arndt-Lappe 2022). Minimal grammatical change defined inside the typological system has been hypothesized to be a characteristic feature of language change (Alber & Meneguzzo 2016; Alber & Kokkelmans 2022; Apostolopoulou 2022), a hypothesis confirmed by the results of the experiment in this study. Name truncation therefore is characterized by structured, rather than random variability: truncation patterns can be ordered by age and conform to an expected path of (minimal) diachronic change.

Keywords: name truncation; prosodic morphology; language change; Optimality Theory; crowdsourcing; predictability; Property Theory; questionnaire study

Resum. *La predictibilitat del truncament de noms tenint en compte el canvi lingüístic*

El truncament de noms propis com en els hipocorístics anglesos (Robert > Rob) s'ha considerat un procés de formació de paraules marginal o extragramatical, a causa de la suposada impredictibilitat i irregularitat dels seus outputs (veg. Dressler & Merlini-Barbaresi 1994, entre d'altres). En aquest estudi, es parteix de la hipòtesi que la coexistència de múltiples patrons de truncament en una sola llengua es deu, en alguns casos, al canvi lingüístic: els patrons més antics conviuen amb els més recents, i això crea una falsa impressió de variabilitat aleatòria. Aquesta hipòtesi es comprova en un experiment en què es demana als participants que triïn el nom escurçat més adequat per a una persona d'un determinat grup d'edat, en un diàleg. L'experiment es va dur a terme a través de la plataforma de proveïment participatiu AlpiLinK (<https://alpilink.it/>) per a parlants de varietats romàniques i germàniques del nord d'Itàlia. Els resultats de l'experiment mostren que els patrons de truncament objecte d'anàlisi per als noms italians i alemanys difereixen significativament pel que fa a l'edat a què els participants els associen. Un segon resultat d'aquest estudi és que els patrons de truncament diacrònicament propers difereixen mínimament en les gramàtiques que corresponen a cada patró en un model formal de truncament implementat com a sistema tipològic en la Teoria de l'Optimitat (Alber & Arndt-Lappe 2022). S'ha hipotetitzat que el canvi gramatical mínim definit dins del sistema tipològic és una característica del canvi lingüístic (Alber & Meneguzzo 2016; Alber & Kokkelmans 2022; Apostolopoulou 2022), una hipòtesi confirmada pels resultats de l'experiment en aquest estudi. Per tant, el truncament de noms es caracteritza per una variabilitat estructurada, més que no pas aleatòria: els patrons de truncament es poden ordenar per edat i segueixen un procés de canvi diacrònic (mínim) esperable.

Paraules clau: truncaments; morfologia prosòdica; canvi lingüístic; Teoria de l'Optimitat; proveïment participatiu; predictibilitat; Teoria de les Propietats; estudi mitjançant qüestionaris

Table of Contents

1. The issue of predictability in name truncation

The truncation of names such as *Robert* > *Rob* in English has been described as a productive process of word formation in many languages over the last decades. Typological surveys furthermore point to a restricted number of shapes that a truncated name can take, clearly definable in terms of the size of the truncatum and of its anchoring properties, i.e., the portion of the base name that it preserves (see Alber & Arndt-Lappe 2012, 2022, 2023, for overviews). Nevertheless, some analysts consider processes of this type not to be part of word formation proper. Ronneberger-Sibold (2015), for instance, classifies name truncation, together with other processes, as word-creation, which she considers outside of the productive models or rules of word formation. One argument in favor of the extragrammaticality (Dressler & Merlini-Barbaresi 1994; Dressler 2000; Mattiello 2013) of name truncation is its presumed unpredictability, linked to the observation that a single language may display several ways to truncate the same name. In Italian, for instance, various patterns of name truncation are attested, which can be classified according to size and anchoring:[1]

(1) Truncation patterns in Italian (Alber 2010; Arndt-Lappe 2018)[2]

Size	**base name – truncatum**	**meaning/function**
1 syllable	Francésca – Fra	hypocoristic
2 syllables	Francésca – Frán.ce	
variable in size	Bárbara – Bá Robérto – Ro.bé Antonélla – An.to.né	Southern/Central Italian vocatives
Anchoring		
Left	Francésca – Fra Francésca – Frán.ce	hypocoristic
Right/Stress	Francésca – Cés.ca	
Left-to-Stress	Antonélla – An.to.né	Southern/Central Italian vocatives

The pattern exemplified by the short name *Fra* for the base name *Francésca* conforms to a monosyllabic template and can be described as left-anchoring, since it copies the left edge of the base name *Francésca.* The pattern *Francésca – Frán.ce* is also left-anchoring, but the truncatum in this case is disyllabic.

1. This table contains only a few of the many truncation patterns attested for Italian (see Alber 2010, for an overview). Base name-truncation pairs are meant to exemplify single patterns, but each of the base names can itself be truncated in many more ways (e.g., *Francésca* > *Fránce, Fra, Francé, Césca, Checca*, etc.).
2. Acute accents indicate main stress in polysyllables; syllables are separated by dots in the truncatum, but not in the base, since the number of syllables of the base is irrelevant for the process of truncation.

Francésca – Cés.ca is an example for a disyllabic pattern which is ambiguous between being classified as right- or as stress-anchored, or even as anchored to both stress and the right edge of the base name. This ambiguity cannot be resolved easily since most Italian names have penultimate stress. *Antonélla – An.to.né* exemplifies a pattern unambiguously anchoring both to the left edge of the base name and to its stressed syllable. Depending on the number of syllables covering that stretch of the word, the size of the truncatum varies. Most of the patterns in (1) have a hypocoristic meaning. However, the pattern represented by *Antonélla – An.to.né* is used exclusively as a vocative (see Alber & Arndt-Lappe 2023 and Section 2 for discussion).

(Tyrolean) German truncated names work in essentially the same way (see, e.g., Arndt-Lappe 2018; Alber & Kokkelmans 2022), except for the fact that a vocative pattern like the Southern/Central Italian one is not found in Tyrol. In an overview like that in Table (1), we could, for example, list the hypocoristic truncations *Kath*, *Kátha*, or *Rína* for the base name *Katharína*.

In this paper, we argue in favor of the process of name truncation being predictable, in the sense that its output does not vary randomly. If this hypothesis is correct, this means for a language like Italian that distinct short name outputs for the same name must differ according to some (socio)linguistic parameter. Even if *Fra, Frán.ce* and *Cés.ca* are all possible short names for *Francésca,* there would then be something which distinguishes them and determines their use. We hypothesize that there are several parameters which distinguish truncation patterns in the same language: (a) patterns may be distinct in meaning and function, (b) patterns may vary in their geographical/dialectal distribution, (c) patterns may differ with respect to their age, i.e., patterns may differ as to when their period of productivity occurred. Examples for each of these parameters will be given, but the main focus of this paper is on (c). Specifically, it will be shown that speakers of Romance and Germanic varieties spoken in Northern Italy distinguish Italian and German truncation patterns with respect to the age of the name bearers to which they are associated. This, we argue, is evidence for the fact that patterns vary with respect to the point in time when they were most productive. A variety of patterns in a single language hence may be the result of language change.

If language change indeed plays a role in distinguishing between older and more recent name truncation patterns, it is possible to construct a second argument in favor of predictability. It has been claimed that especially in the domain of prosodic morphology language change proceeds in minimal steps, affecting the grammatical architecture of the language one small step at a time (Alber & Meneguzzo 2016; Alber & Kokkelmans 2022; Apostolopoulou 2022). Minimal grammatical change is defined in these proposals as a minimal change in the ranking conditions defining a grammar, i.e., in its typological properties (see Alber & Prince 2015, 2021 for Property Theory in Optimality Theory). If this approach to language change is on the right track, it predicts that not all forms of language change are expected to occur with the same degree of probability. If, in a language, we find a series of name truncation patterns and can order them according to their age, this order should reflect the changes predicted to be minimal in a formal typol-

ogy of truncation. In the second part of the paper, we compare the diachrony of name truncation patterns as emerging from our empirical study with the predictions made by a formal model of truncation (Alber & Arndt-Lappe 2022). We conclude that our empirical findings are interpretable as involving for each change in time a minimal change in the ranking conditions defining the patterns. At least where language change is involved, multiple truncation patterns in the same language are therefore predictable also in this sense: they should differ minimally in their grammar, if they differ minimally in their age.

The paper is structured as follows: in Section 2, we give examples of differences in meaning/function, geographical/dialectal distribution and perceived age of name truncation patterns, drawn from previous work; in Section 3, we show how the age of truncation patterns can be investigated and report the results of our experiment for Italian and German names in Northern Italy; in Section 4, we compare our empirical results to the predictions made by a formal model of truncation; Section 5 contains the overall conclusions.

2. Predictability of name truncation: The role of meaning and function, space, and time

Besides supposed variability as that encountered in Italian, a further feature of name truncations that has contributed to their characterization as extragrammatical is the fact that their meaning or function remains somewhat elusive and understudied (see Alber & Arndt-Lappe 2023 for an overview; Schneider 2003, for English, is an exception). The meaning of truncated names is close, in many cases, to that of the output of processes of evaluative morphology such as diminutives. As in diminutive formation, the referent of the base of derivation can be the same as that of the derived form: *Francesca* and *France* can refer to the same person. In the case of hypocoristic formation, the difference between the two forms can be characterized in terms of a higher degree of familiarity with the referent (or smaller social distance) expressed by the truncated form, compared to the social relation expressed by the base name. It seems to be clear, in any case, that the meaning of a hypocoristic depends on the contexts in which they are used and therefore is determined by pragmatic factors (Dressler & Merlini Barbaresi 1994; Schneider 2003; Merlini Barbaresi & Dressler 2020).

For the purpose of this paper, it is important to notice that even though it is difficult to pin down the exact meaning of a name truncation pattern, there are cases where patterns can be distinguished on the basis of their meaning or function. For some patterns, the difference between them lies in the degree of familiarity expressed by the truncated name. Thus, in English, the use of the short name *Ed* for *Edward* can be appropriate in contexts where the use of *Eddie* is not, since *Eddie* expresses a higher degree of familiarity than *Ed* (Schneider 2003: 145-149; Alber & Arndt-Lappe 2012, 2023). Monosyllabic truncations like *Ed* thus express greater social distance than *i*-truncations like *Eddie.* Other truncation patterns are linked to a very specific meaning or function, e.g., that of a vocative. Thus, while short names in Italian can, in general, be used both as proper nouns with a hypocoristic meaning as well as voca-

tives, the pattern exemplified by *Antonélla-An.to.né* can be used only as a vocative.[3] The sentence *Antoné, spicciati!* 'Antoné, hurry up' is therefore grammatical, while a sentence in which the same short name is used without a vocative meaning is not, cf. **Oggi ho visto Antoné in città* 'Today I saw Antoné in town' (Alber & Arndt-Lappe 2023; a similar restriction to a vocative meaning occurs in Georgian truncated kinship terms, see Abuladze & Ludden 2013: 31). In this sense, the patterns represented by *Francésca-Frán.ce* and *Antonélla-An.to.né are* predictable. The former are used both with a non-vocative and a vocative meaning, while the latter are limited to vocative contexts.

Dialectal variation plays a role as well in distinguishing patterns from each other. No detailed geolinguistic studies exist on the distribution of name truncation patterns in space, but studies of single patterns often mention their regional distribution. For example, the doubly-anchored pattern exemplified by *Antonélla-An.to.né*, which varies in the size of the truncatum, is attested in Central and Southern, but not in Northern varieties of Italian (Alber 2010; D'Alessandro & van Oostendorp 2016; Kenstowicz 2019; see also Vanrell & Cabré 2011, for Catalan varieties of Sardinia exhibiting this pattern). Also in this sense, some patterns are predictable because they occur only in certain regional dialects.[4]

Besides variation in space, name truncation is also characterized by variation in time. For example, Germain (2014: 142-147) mentions, with respect to truncated names in Wallonia, Belgium, that there has been a first wave from the Middle Ages to the 16th century during which right-anchored truncations were especially productive (e.g., *Colas* from *Nicolas*). People continued to use this type of truncated names until World War II, when a second wave started to spread and made left-anchored truncations popular (e.g., *Nico* from *Nicolas*). The simultaneous presence of patterns like *Colas* and *Nico* in modern French is therefore not due to random variation but more likely to language change.

For Medieval Tuscan Italian, instances of a name truncation pattern are reported which can be described as anchoring to the left and to the right edge of the base name, as well as to the stressed syllable, with the possibility to syncopate medial segments in order to satisfy a disyllabic template. Examples of names truncated this way are *Duránte-Dán.te, Beatríce-Bí.ce, Benedétto-Bét.to, Battísta-Bís.ta, Federígo-Fí.go, Giovánni-Gián.ni* (Brattö 1953: 44; Scherillo 1896: 45f.). In Modern Italian, only the short names *Bice* and *Gianni* have survived as representatives of this pattern. They coexist with other short names in Italian but are clearly distinct from the modern patterns with respect to the period of time in which they were productive.[5]

3. The claim that this pattern is used exclusively as a vocative is based on the elicitation of judgements by speakers of Rome. We cannot be sure that *Antonélla-An.to.né* is used exclusively as a vocative in all regions where the pattern is attested.
4. In his historical grammar of Italian dialects, Rohlfs observes in the chapter on "Koseformen von Vornamen" ('hypocoristics of first names') that Italian hypocoristics "[…] sind regional zum Teil stark verschieden." ('In part they differ strongly from region to region'; Rohlfs 1949: 513, §319). For many of the patterns that he discusses he mentions a specific region where the form is most commonly used.
5. Scherillo (1896) discusses in detail the multitude of short name patterns *en vogue* during Dante's lifetime. He mentions another reason which according to him led to the coexistence of multiple

That language change does play a role in the simultaneous presence of multiple name truncation patterns in a single language has been shown recently in a study by Alber & Kokkelmans (2022). In their study, they test six French and seven German name truncation patterns as to the perceived age that speakers associate with each pattern (see also Boschiroli 2017 for a similar study for Italian). In the experiment, speakers are asked how old a person bearing the short name X for the base name Y might be. For both languages a relative chronology of older and younger truncation patterns can thus be established. Alber & Kokkelmans (2022) interpret the simultaneous presence of older and younger patterns in the same language as the result of a relatively fast turnover of truncation patterns in the speech community. As older patterns start to sound obsolete, new patterns emerge. If the presence of several patterns in the same language is (at least to some extent) the result of language change, we have identified one more factor distinguishing name truncation patterns from each other.

In sum, it is far from clear that the fact that several short names are attested for a single base name in some language can count as an argument for the unpredictability of the process of name truncation, in general. *Frán.ce* is different from *Cés.ca*, because, as will be shown below, it is associated with younger name bearers than *Cés.ca. An.to.né* and *Án.to* are both possible short names for *Antonélla*, but the former is restricted in its use to the Southern and Central varieties of Italian and, furthermore, it can be used only as a vocative, while *Án.to* is not limited to vocative contexts.

In the remainder of this paper, we concentrate on the factor of language change as a major influence on the emergence of different truncation patterns.

3. Name truncation in the Romance and Germanic varieties of Northern Italy

3.1. Setting up the experiment

Data to test the perceived age of truncation patterns was collected with the help of a crowdsourcing platform implemented in the context of the research project AlpiLinK (<https://alpilink.it/>; Rabanus et al. 2024). AlpiLinK aims to collect data for all Romance, Germanic and Slavic varieties of Northern Italy. The target varieties include non-standard dialects such as the Germanic cluster of Tyrolean varieties, as well as recognized minority languages like (Romance) Ladin or (Germanic) Mòcheno. Over the AlpiLinK platform we offered two distinct questionnaires, one with Italian names for the speakers of Romance varieties and one with German names for the speakers of Germanic varieties. It is worth noting that speakers of Germanic minority languages are typically L2 speakers of (Standard) Italian, while Romance speakers more rarely speak a German variety.

Four truncation patterns were selected for the questionnaires. The patterns were chosen according to their hypothesized age, based on what could be gleaned

patterns: since it was customary to name children after their father or mother, it was necessary to distinguish them from their parents by using a truncated or suffixed name (Scherillo 1896: 49, fn.2).

from previous studies (Arndt-Lappe 2018; Alber & Kokkelmans 2022 for German; Alber 2010; Boschiroli 2017 for Italian). To obtain a solid basis of comparison, we chose truncated names for each pattern that were shown in previous studies to be frequently used, and we consistently favored truncated names that were derived from the base name with the smallest change in segments (e.g., *Césca* but not *Chécca*, for the base name *Francésca*). The patterns are illustrated in (2) for the Italian base name *Francésca.* The same patterns were tested both for Italian and for German names, since all four of them are attested in both languages.

(2) Tested truncation patterns – exemplified for the base name *Francesca*

	base name	**truncated name**	**size of the template**	**anchoring**	**hypothesized perceived age**
A	Francesca	Fra	monosyllabic	left	young ↓ old
B		Frán.ce	disyllabic	left	
C		Frán.c-i	disyllabic, *i*-suffixed	left	
D		Cés.ca	disyllabic	right/stress	

Based on the results of previous studies of Italian and German (Boschiroli 2017; Alber & Kokkelmans 2022), we hypothesized that Pattern A (monosyllabic, left-anchoring) would be associated by speakers to young name bearers, Pattern B (disyllabic, left-anchoring) to relatively older names bearers, the *i*-suffixed Pattern C to yet older ones, and D, which is a disyllabic pattern ambiguous between right- and stress-anchoring, should obtain the highest values for perceived age.

To test this hypothesis, three Italian and three German base names were chosen, which each had attestations for at least three short names classifiable for the patterns A-D, above.

The selected base names were controlled for frequency and average age of the name bearers. This was necessary since we wanted to choose base names that are well-known (hence of high frequency), but that are at the same time not associated with very young or very old name bearers. For instance, the name *Josef* is the most frequent German male name in the statistic surveys available for the region of South Tyrol (ASTAT 2018), but at the same time the average age of men bearing this name is 64. This means that *Josef*, although well-known as a name, is probably associated with older name bearers and this association might bias the perceived age that speakers assign to its short names.

Base names were chosen that were among the 20 most frequent Italian and German names in the region. As a reference we used the statistical survey of the city of Trento (#TrentoInCifre 2023) and the statistical survey of the Autonomous Province of Bozen-Bolzano, South Tyrol (ASTAT 2018) for Italian and German, respectively.[6] The average age of the name bearers of the selected names was

6. These were the most detailed statistical surveys available for Italian and German names in Northern Italy, and the only ones offering also information about the average age of name bearers (see also Raus 2024). Unfortunately, no data is available as to the development of the frequency of names

between 27 (*Matthias*) and 48.2 (*Francesca*). In addition, we controlled for potential frequency effects among the alternative truncated names presented in the study, making sure that all of them were reasonably common in the language concerned. The selection criteria just outlined turned out to be quite difficult to fulfill in some instances, leading to some minor imperfections in the sample.[7] The following table contains all base names and the truncated names associated with the four patterns, in the orthographic representation that was used for the study:

(3) Base names in the experiment and their truncations

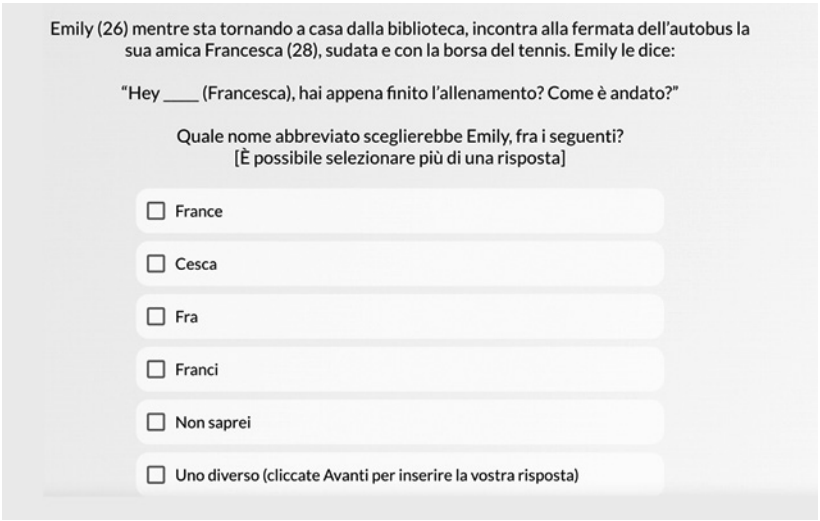
	base name	**A**	**B**	**C**	**D**
Italian	Francesca	Fra	France	Franci	Cesca
	Giovanni	Giò	Giova	Giovi	Vanni
	Matteo		Matte	Matti	Teo
German	Matthias	Mats	Matthi		Hias
	Theresa	Tess	There	Theri	Resi
	Isabella		Isa	Isi	Bella

Participants were asked to choose the short names they thought were appropriate for the name of one of two personae mentioned in a dialogue that was presented to them in written form. The age of the personae in the dialogue varied according to three age groups (18-30, 30-60, > 60). Both personae in each dialogue were assigned ages of the same age group to control for potential effects caused by politeness between age groups. The dialogues contained similar speech acts: one of the two personae was addressed with an offer or a question (cf. Schneider 2003 on effects of speech act on the use of diminutive and hypocoristic forms). Call-vocatives were avoided, but normal vocatives were allowed (see (4) below). The dialogues centered around similar topics, from the domain of 'family and friends'. In a forced-choice design, the task of the informants was to choose one or more appropriate short names for the dialogue; they also had the possibility to choose "I don't know", and there was a field under 'other' in which they could insert their own suggestions for a short name appropriate for the context. An example of the task for the Italian base name *Francesca* is given in (4).

over time. As noted by Jennifer L. Smith (p.c.), the best base names would be those without spikes in their frequency over time, since they would most probably least influence the perceived age of the truncated names.

7. The German base name *Isabella* is not among the 20 most frequent names according to the statistical survey, yet, we think, rather well-known; the short name *Resi* for *Theresa* is somewhat different from the other short names representing Pattern D since, though stress-anchoring, it is suffixed with *-i*; the short name *Matthi* for *Matthias* appeared in the task but had to be excluded from part of the analysis, since it is ambiguous between patterns B and C.

(4) The choice of a truncated name - task presented to participants[8]



All participants were presented with the 9 stimuli in the same order. For German, this was Theresa – Isabella – Matthias, repeated three times in this order, and for Italian it was Francesca – Matteo – Giovanni, again repeated three times in this order. For both languages, the sequence of the age groups of the personae was: young – old – mid – old – mid – young – mid – young – old. This fixed order ensured that no two identical names or age groups would follow one another in the stimulus sequence. The order of the short names listed as possible answers, however, was randomized for each stimulus.

The data was collected in the context of the AlpiLinK data campaign of 2023-2024 among speakers of all varieties investigated by the AlpiLinK project. The AlpiLinK questionnaires also provided sociolinguistic information about the participants such as age, gender, place of origin, and the linguistic variety that participants chose as their native variety for the tasks.

A decisive role for the success of the data campaign was played by the citizen science project VinKiamo (<https://alpilink.it/en/vinkiamo/>), for which high school students helped older speakers to submit their data to the crowdsourcing platform, thus contributing to create a database balanced for age of the participants. The age of participants ranged from 10 to 101, with an average of 48.1. For our analysis, we were able to elicit the data of 163 speakers of a Germanic variety (all

8. Translation of the text: "Emily (26) encounters her friend Francesca (28) at the bus stop while she is returning home from the library. Francesca is sweating and carries tennis equipment. Emily says to her: 'Hey ___ (Francesca), did you just finish your training session? How did it go?'. Which of the abbreviated names would Emily choose, among the following? It is possible to select more than one answer. *France, Cesca, Fra, Franci*, *I wouldn't know*, *A different one* (click on 'Forward' to insert your answer)." The full set of experimental scenarios is available at <https://osf.io/3j694>.

but one native speakers of a Tyrolean dialect) and of 582 speakers of a Romance variety. Short names suggested by the participants (in the open field 'A different one') were assigned to one of the four patterns, when possible (e.g., *Checca* for *Francesca* was assigned to Pattern D). Interestingly, for some names single participants suggested short names fitting one of the four patterns but which we had not listed among the options because we were not aware of their existence. This was the case of *Mat* for *Matteo* (Italian speakers, a possible instance of Pattern A) and *Mátte* for *Matthias* (German, a possible instance of Pattern B). These names were not integrated in the analysis, since it is impossible to know how the overall group of participants would have rated them.

The analysis was performed on version 1.1.1 of the AlpiLinK corpus (Rabanus et al. 2024). The overall number of collected items (valid answers) used for the analysis was 5513 (this number does not include user-suggested truncations nor the ambiguous truncation *Matthias – Matthi*, which can be classified as Pattern B or Pattern C).

3.2. Analysis

A multinomial logistic regression analysis was carried out to test which variables co-determine participants' choices among the four truncation patterns presented in the experiment (implemented as the *multinom()* function in the *nnet* package, Venables & Ripley 2002, in R, R Core Team 2023). The statistical analysis includes the truncation pattern (A, B, C, D) as the dependent variable (PatternsABCD) and tests for the following fixed effects: the age group of the persona in the dialogue (NameAgeSetting), the gender of the base name in the dialogue (NameGender), the age and gender of the participant (OwnAge, OwnGender), and whether the base name is Italian or German (GermanicRomance). In addition, we include information about the place of origin of the (linguistic variety of the) participants, coded in terms of longitude and latitude.

Recall from Section 3.1 that in the questionnaire we used existing truncated forms for our six base names. This has several consequences for our analysis. First, we were not able to include all four patterns (A, B, C, D) for each base name in the questionnaire. Thus, Pattern A (monosyllabic, anchoring to the beginning of the base name) was not represented among the options for the names *Isabella* (German) and *Matteo* (Italian), and patterns B (disyllabic) and C (disyllabic, with a final suffix *-i*) were homophonous for *Matthias* (German), and hence had to be excluded from the main part of the statistical analysis. What is more, we expect that the frequency of use and the level of entrenchment of a truncated form will have an influence on participants' choices. Such frequency information is not readily available for truncated forms in standard corpora because most of these forms belong to an informal register of language use, and because such registers are traditionally strongly underrepresented in existing large language corpora. We are therefore not able to include the frequency of the truncated form as a covariate in our statistical analysis. In order to do justice to the heterogenous nature of our dataset thus described, our analysis will proceed in three steps.

To test our hypothesis about the temporal dynamics of truncation patterns (cf. esp. the table in (2) above), we will first present separate multinomial models for each of the base names in the study and investigate to what extent participant choices of truncation patterns are correlated with dependent variables encoding temporal dynamics. These are (a) the age of the name bearer (NameAgeSetting, levels: 'young', 'mid', 'old'), and the sociodemographic variable measuring the age of the participant (OwnAge, in years). In a second step, we will compare selectional preferences across different base names on the basis of a subset of the data for which there is no missing information. This will enable us to go beyond the variables capturing temporal dynamics and to factor in potential differences relating to the language concerned (GermanicRomance, levels: 'Italian', 'German') and semantic differences between truncation patterns, captured here in terms of the gender of the name bearer (NameGender, levels: 'female', 'male'). As a last step, we will zoom in on the Italian subset of our data and incorporate potential regional predictors, investigating to what extent the selectional patterns of our participants are also co-determined by the region of residence.

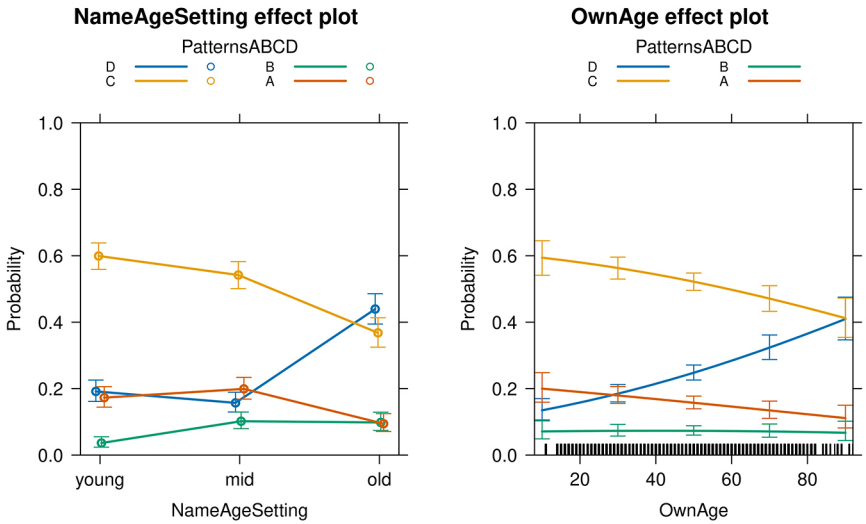
All findings will be presented by means of effects plots, and statements about the significance of effects will be based on z-tests. A full documentation of the summary statistics of all models presented can be found in the appendix. All data and R scripts are available at <https://osf.io/3j694/>.

The figure in (5) displays the results of the multinomial models run on each individual base name in the Italian subset of our data. Only NameAgeSetting and OwnAge are included as predictors. Note that the effects reported here remained stable also when we included more predictors coding for sociodemographic information about the respondents of the survey.

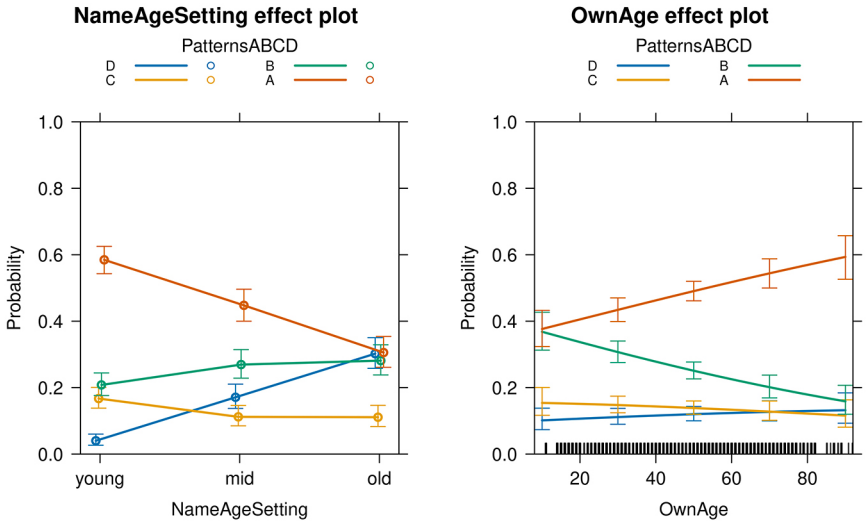
In the figure in (5), the y-axis corresponds to the predicted probability with which a pattern has been selected by the participants, proportionally to the probability of other patterns. This means that adding up all probabilities vertically in the figure (e.g., for all probabilities where *Francesca* is young: {A: 101 = 17%; B: 21 = 4%; C: 348 = 59%; D: 116 = 20%}), one obtains a total probability of 1 (e.g., 586 answers = 100%). The same holds for the right-hand part of the figures, where the probability of a participant selecting an A-, B-, C-, or D-pattern increases or decreases relatively to the probabilities of the other patterns as a function of the age of the participant.

(5) Individual models, Italian names. Y axes show the predicted probability of the respective pattern. Reference levels for categorical variables: 'D' (PatternsABCD), 'young' (NameAgeSetting).

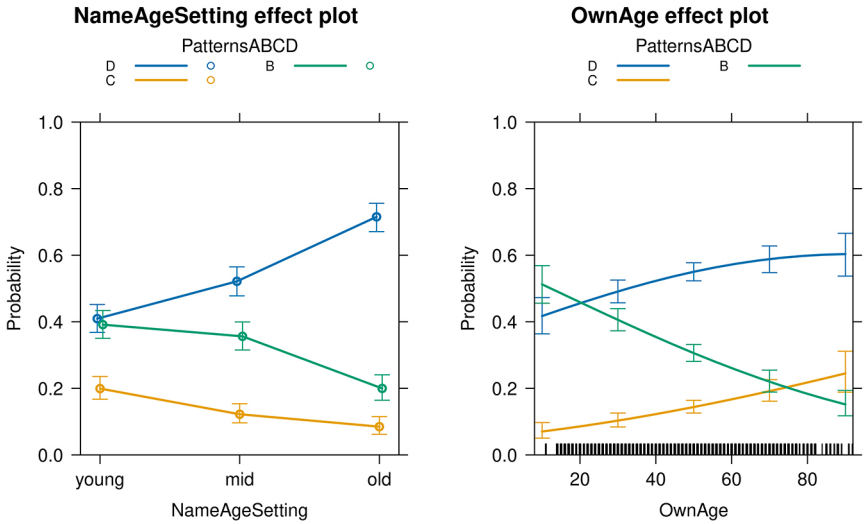
a. Francesca (N = 1625)



b. Giovanni (N = 1348)



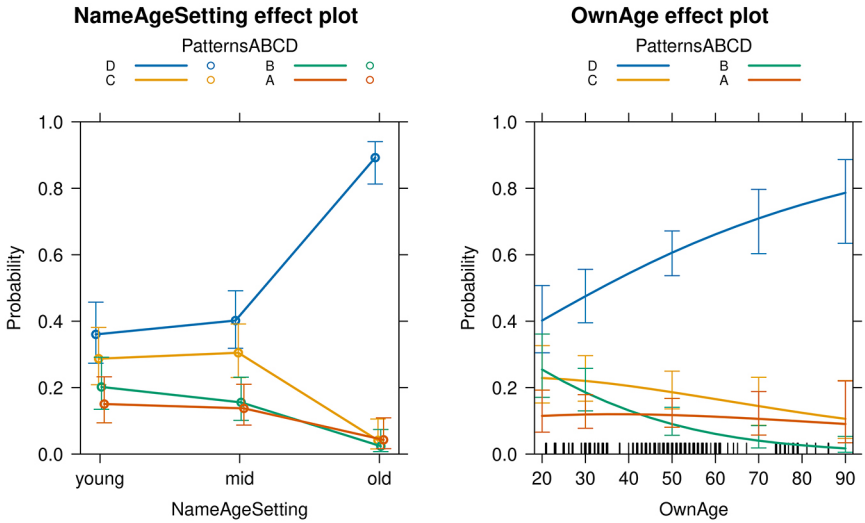
c. Matteo (N = 1491)



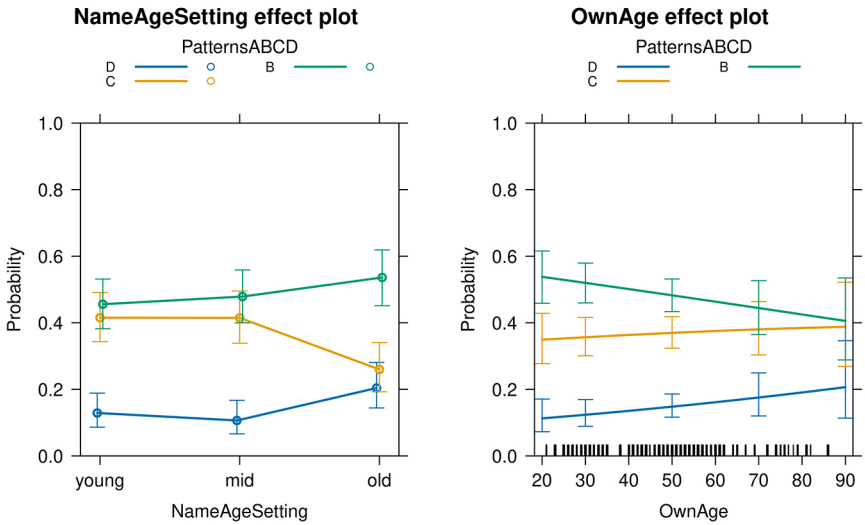
Even though all base names are indeed different, still very similar selection patterns are discernible across names. For *Giovanni* and *Francesca*, we have data on participants' preferences for all four patterns (A, B, C, D); for *Matteo*, no data are available for Pattern A. All three names show a significant effect of age group of the name bearer (NameAgeSetting), in that Pattern D is more commonly selected for 'old' name bearers than for younger name bearers. The difference between 'young' and 'mid' aged name bearers is significant for *Giovanni* and *Matteo*, but not for *Francesca*. Pattern A, if present, shows the opposite trend of Pattern D, being more dispreferred for old name bearers than for name bearers classified as 'young' (*Giovanni, Francesca*) and 'mid' (*Giovanni*). Patterns B and C show less clear-cut age grading. We will see below that one reason is that selection of these patterns is co-determined by other factors as well that are not part of the current model. The effects of the variable OwnAge, which codes the age of the respondent, do not always mirror the effects of NameAgeSetting. Whereas effects for *Francesca* are clearly parallel in that older respondents prefer Pattern D while younger ones have a preference for Pattern A, for *Giovanni* and *Matteo* only a partial correspondence can be observed with regard to patterns A and D. Also, the selection of Pattern B is negatively correlated with the respondent's age for *Giovanni* and *Matteo*, while for *Francesca* this pattern is a clear minority pattern across ages. The figure in (6) now presents the parallel individual models for the German names.

(6) Individual models, German names. Y axes show the predicted probability of the respective pattern. Reference levels for categorical variables: 'D' (PatternsABCD), 'young' (NameAgeSetting). The figure for the name *Matthias* (6c) visualizes the effect of a simple logistic regression model, as only two options for truncated forms were available for this name (A, D).

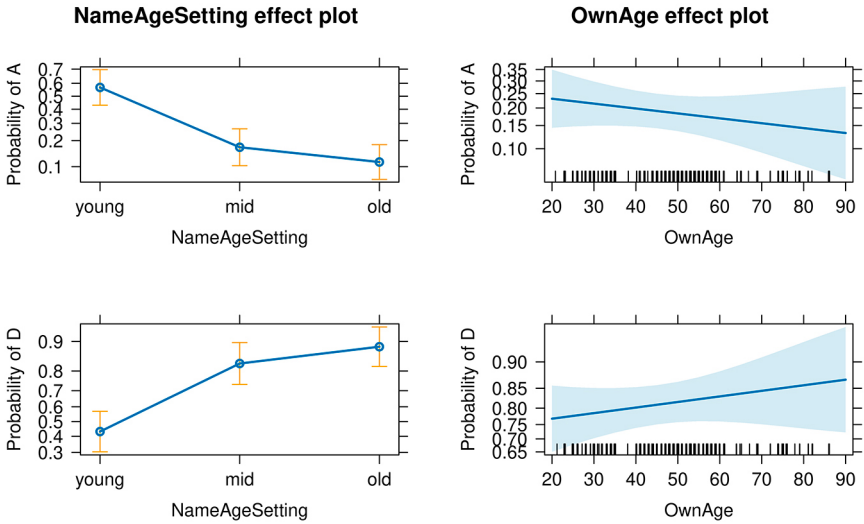
a. Theresa (N = 331)



b. Isabella (N = 454)



c. Matthias (N = 264), A-pattern (above), and D (below)



The models show very similar effects to those observed for the Italian names. Note, however, that overall models are based on a much smaller number of observations than the analysis of the Italian data. Individual names differ considerably in the degree to which a pattern is selected. For example, the predicted probability of Pattern D in the age group 'old' is 90.1% for the name *Theresa*, but only 21.8% for the name *Isabella*, in spite of the fact that there were more options available for *Theresa* in the questionnaire ('A', 'B', 'C', 'D') than for *Isabella* ('B', 'C', 'D'). Still, for all three names in the dataset, Pattern D is consistently selected more often for name bearers in the 'old' age group than for name bearers in the 'young' and 'mid' age groups. Pattern A also shows a similar behavior as in the Italian dataset. For both *Theresa* and *Matthias*, it is selected significantly less often for 'old' referents than for 'young' (*Theresa*, *Matthias*) and 'mid' (*Matthias*) referents. Pattern C shows a parallel, age-graded trend like Pattern A in the German data (for *Isabella* and *Theresa*), in that it is selected significantly less often for 'old' name bearers than for 'young' and 'mid' referents. With regard to Pattern B, no consistent trend can be observed for *Theresa* and *Isabella*. Like in the Italian data, the trends observed in the variable NameAgeSetting are not exactly parallel to those observed in the variable encoding the age of the respondents (OwnAge). One generalization shared among all three names, however, is that the probability that Pattern D is selected increases with the age of the respondent. Note however, that this correlation does not reach statistical significance for *Matthias* and *Isabella*.

In general, we see that consistent age grading effects of truncation patterns are more clearly associated with the name bearers than with the age of the respondents. We interpret this observation as reflecting the fact that both older and younger speakers associate truncation patterns with age in a similar way. This is not too

surprising, since both older and younger speakers are probably exposed to reference to people of all age groups by their truncated names.

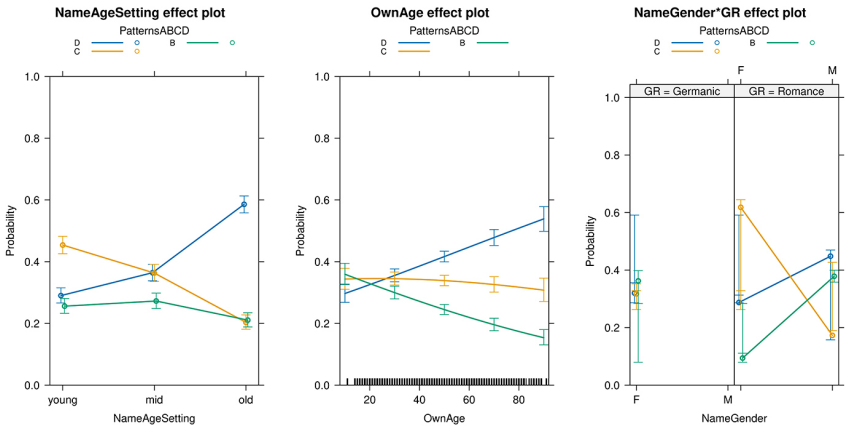
We do, however, find some usage preferences that depend on the respondents' age, and that are independent of the perceived age of the pattern. These are inconsistent across base names and truncation patterns, and can therefore not be explained by the age of the truncation pattern.

Summing up our findings so far, we saw that across all names investigated and across the two languages, Pattern D is clearly the oldest pattern and Pattern A is the youngest. This is evidenced by the consistent age grading patterns they display with regard to the age of the name bearers. Patterns B and C, by contrast, do not show consistent age-related effects. This raises the question whether selection of these two patterns in particular is subject to further constraints that have not been captured in the variables measuring temporal dynamics.[9] Furthermore, our analysis so far has not been able to test for effects that might hold across individual names. To address these issues, we will now present an analysis that includes data from different base names, to allow us to make the comparison. To avoid problems connected to missing information in some data cells and in order to maximize the information available to our model, we create a subset that includes only the patterns B, C, and D (as 'A' information is missing for two names). In addition, we exclude data for the base name *Matthias* (German), as we have no information about patterns B and C for this name. The subset thus created hence includes the data for participants' selection of B, C, and D for the base names *Francesca*, *Giovanni*, *Matteo* (Italian), and *Isabella* and *Theresa* (German). The overall number of observations in our subset is N = 4330. Using again a multinomial regression analysis, we now also include a variable distinguishing Italian and German names (GermanicRomance) and a variable coding for the gender of the name bearer (NameGender). The latter may capture a potential source of semantic differences between truncation patterns: gender (cf. Schneider 2003 on gender-related functional differences among truncation patterns in English). We ran several alternative models which tested for all possible effects of GermanicRomance; these involved testing GermanicRomance as an individual predictor as well as testing interactions of GermanicRomance with all other predictors. Model comparison revealed that GermanicRomance is confounded with gender in our dataset, in the sense that the main observable difference between German and Italian names is that Italian shows a stronger tendency to use Pattern C for female name bearers in our dataset than German. Since, however, our dataset does not contain information on German male name bearers, the variable GermanicRomance emerged as significant when used as an individual predictor as well as when used in interaction, with no difference between the two statistical models in terms of their goodness-of-fit. Crucially, models testing for potential interactions of the variable

9. Regarding the change in truncation preferences from D to A, one reviewer points out that it could have various motivations, e.g., enhancement of lexical access via the left edge, or the result of a constraint like StrongStart (Selkirk 2011). We leave this question open for further research and concentrate here on the spread of this change.

GermanicRomance with the age-related predictors (NameAgeSetting, OwnAge) showed that the two languages essentially follow the same trends, so that no clear differences can be found between them.[10] The figure in (7) displays the results of the most parsimonious of our models, in which an interaction of NameGender and GermanicRomance is included, allowing us to test if there is a gender effect only for the Italian names.

(7) Model predicting the pattern selected ('B', 'C', or 'D') for the names *Francesca*, *Giovanni*, *Matteo* (Italian), and *Isabella*, *Theresa* (German). Y axes show the predicted probability of the respective pattern. Reference levels for categorical variables: 'D' (PatternsABCD), 'young' (NameAgeSetting), 'Germanic' (GermanicRomance), 'F' (NameGender). Note that 'GermanicRomance' has been abbreviated to 'GR' on this figure for reasons of space.



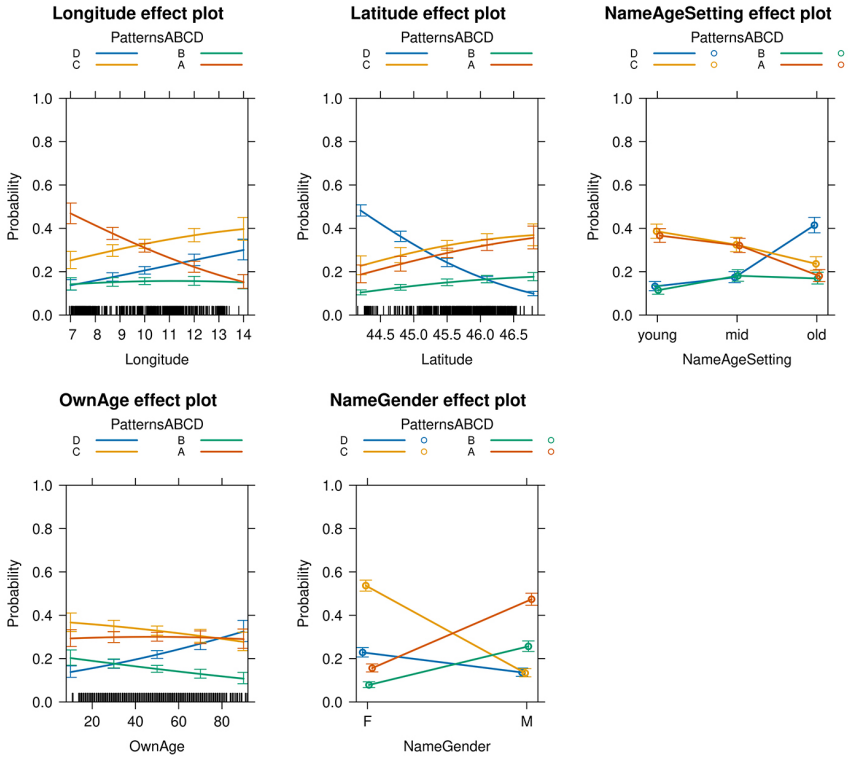
For the age-related variables (NameAgeSetting and OwnAge), the model in (7) shows the effect for Pattern D that we have already established in the models on individual names from both languages. In addition, there is a significant correlation of NameAgeSetting and the predicted probability of Pattern C, in that Patterns C is less likely to be selected for 'old' name bearers as compared to 'young' and 'mid aged' referents. A similar, but much less strong trend can be observed for Pattern B. For Pattern B, but not for Pattern C, this trend is also reflected in the

10. No significant interaction was found for GermanicRomance and OwnAge, whereas in a model including an interaction of GermanicRomance and NameAgeSetting the interaction did emerge as significant for Pattern C (at the level of $0.05 > p > 0.01$). Thus, Pattern C shows a significant age grading effect in Italian (predicted probability of C is positively correlated with the age of the name bearer: 'young' > 'mid' > 'old'). In German we observe a similar, but much weaker trend, which does not reach statistical significance. A similar, non-significant trend is also found for male name bearers in Italian when we include a triple interaction of GermanicRomance, NameGender, and NameAgeSetting in the model. We tentatively conclude that there might be a general age grading tendency in Pattern C, but have to leave it to future research to investigate if and in case how far this effect, if substantiated, is modulated by gender and language differences.

OwnAge variable. Whereas these findings do suggest that Patterns B and C are younger than Pattern D, the third panel in (7) clearly shows that Pattern C is also functionally different from Patterns B and D. The predicted probability of Pattern C clearly depends on the gender of the referent of the name. For the Italian names, it is 61.2% if the name bearer is female, and 17.0% if the name bearer is male. For German names, which in this dataset are all female, the mean predicted probability of Pattern C is 30.2%, hence significantly lower than for the Italian female names. Since the absence of male names in the German dataset does not allow us to compare gender effects in German and Italian, and since we saw in our previous analysis of individual names that there are considerable differences between individual names, we have to leave the details of the observed gender effect in German names to future research. Still, it is interesting to note at this point that the gender effect for Pattern C, which we clearly found for Italian, has parallels in the (scant) literature on the functions of truncation patterns. Pattern C comprises disyllabic truncated forms that end in the suffix *-i* (e.g., Italian *Francésca – Frán.c-i*). Analyzing a parallel process in English (e.g., *Margaret – Maggie*) as an instance of diminutive formation, Schneider (2003) observes that such forms are commonly used for female or child referents. We also see gender-related effects for Patterns B and D in the Italian names in (7), which are less strong and go in the opposite direction, as these patterns are more likely to be selected for male than for female referents. Hence, it seems like avoidance of Pattern C for male referents is complemented by corresponding increased preferences of alternative patterns for male names.

We now finally turn to the question of whether there are regional patterns in the participants' preferences. This question can be addressed only on the basis of the Italian data, as these cover a larger region, for which we can assume a dialect continuum. The underlying hypothesis is that more recent patterns should be more commonly selected in the West of the survey region, as compared to older patterns, since in the West we find big urban centers such as Milan and Turin, which might function as the origin of innovative patterns or as a hotspot for the spread of innovative patterns. This is why we are particularly interested in selection preferences for Pattern A, the youngest pattern. We therefore select all data for the Italian base names *Francesca* and *Giovanni*, i.e., those Italian base names for which participants selected among all patterns (A, B, C, and D), to test our hypothesis. This subset of the data comprises 2,973 observations. The figure in (8) illustrates the results of the statistical analysis. The multinomial model includes the variables Longitude and Latitude to refer to the place of origin of participants. In addition, we also included the variables that were found to co-determine the participants' selection of patterns above, both those related to temporal dynamics and those related to gender of the name bearer. The dependent variable is again the pattern selected.

(8) Model predicting the pattern selected ('A', 'B', 'C', or 'D') for the names *Francesca* and *Giovanni* (Italian). Y axes show the predicted probability of the respective pattern. Reference levels for categorical variables: 'D' (PatternsABCD), 'young' (NameAgeSetting), 'Germanic' (GermanicRomance), 'F' (NameGender).



The figure shows first of all that all effects that we have observed in our previous analysis of different subsets (NameAgeSetting, OwnAge, and NameGender) also appear in the current model, providing confirmation for the robustness of the patterns in the data. In addition, we indeed see significant regional differences, in particular between the recent pattern A and the oldest pattern D. The trends for A and D go in opposite directions. Pattern A is more common in locations with lower longitude values (i.e., in the West) than in locations with higher values (i.e., in the East). The opposite is true for the older Pattern D. Latitude effects suggest that the recent pattern A is also more common in the North of the survey region than in the South, and again the opposite is true for the older pattern D. Pattern C also shows significant effects for Longitude and Latitude values, which are, however, somewhat more difficult to interpret. We may tentatively interpret the West-East difference as the spreading of newer patterns from the West, i.e., from metropolitan hubs such as Milan and Turin. The North-South difference could be interpreted

as more conservative patterns being used more frequently in the regions closer to the La Spezia-Rimini isogloss, which constitutes an important separation between Western and Eastern Romance varieties. In alternative, the effects emerging from the analysis could be seen as spreading of innovative patterns from the Northwest to the Southeast. It is beyond the scope of this paper to follow these lines of investigation further. We conclude that the language change hypothesized here is expected to have variation in space as a consequence, since the spreading of a change in linguistic structures usually does lead to regional variation, when patterns spread from centers of innovation to the periphery.

3.3. Discussion of results

In sum, the analysis in the present section has revealed robust evidence for an age-related ordering of three of the truncation patterns investigated, A, B, and D. Participants thus distinguish name truncation patterns into patterns used for younger people and patterns used for older people. We also found that Pattern C is more difficult to place, and we saw that this is likely due to the fact that Pattern C is also functionally, i.e., semantically different from the other patterns. The age of respondents, on the other hand, did not consistently correlate with the choice of truncation patterns.

The table in (9) schematizes the ordering we found. Based on the fact that Pattern C showed somewhat stronger age-grading effects than Pattern B, which were also in line with the behavior of the clearly recent Pattern A, we tentatively group Pattern C with Pattern A in (9), but acknowledge that the placement is tentative, and more research is needed exploring the interaction of function and age for Pattern C.

(9) Relative order of patterns with respect to the age of the personae

	A, C (?)	B	D
Francésca	*Fra, Frán.c-i*	*Frán.ce*	*Cés.ca*
	young ⟶ old		

As expected, Pattern D was associated with personae in the dialogues that belonged to the older age groups. The next pattern on the scale was B. Patterns A and (possibly) C were in turn associated with younger personae. The relative diachrony of the four tested patterns is therefore very close to what we have expected at the beginning. The initial hypothesis of the order A < B < C < D was in part disconfirmed, as we saw that Pattern C does show age-related effects similar to Pattern A, but is subject to additional, functionally motivated constraints.

In addition, we found regional effects for a subset of the data which provide additional support for language change in general and for the relative order proposed in (9), under the assumption that innovative patterns like A originate in big metropolitan centers such as Milan and Turin, which are located in the West. In sum, then, we interpret our findings regarding age-related and regional effects on

participants' selection patterns as evidence that the different truncation patterns attested for Italian and German names are the result of language change: some have emerged recently, hence are perceived as appropriate for younger people, some are of older formation and hence are associated with older name bearers.

4. Connection to a formal model of truncation

Alber & (2016), Alber & Kokkelmans (2022) and Apostolopoulou (2022) pursue the idea that one of the factors influencing the change of linguistic structures over time is minimality of the observed change, in the grammatical sense. They claim that there are cases of language change for which it can be observed that the architecture of the grammar at a given point t_1 in time is minimally different from the grammar of the same language at a later point t_2, i.e., where language change proceeds in minimal intensional steps. They propose furthermore that minimal differences between grammars can be understood as minimal differences between ranking conditions inside the factorial typology generated in an optimality-theoretic system (Prince & Smolensky 2004 [1993]). More precisely, two grammars differ minimally, if they differ in a single *property value*, where typological properties are understood as the defining ranking conditions of a typology (Alber & Prince 2021).

In this section, we consider whether the name truncation patterns appearing in the present study as minimally different in their perceived age can also be understood as minimally different with respect to the property values defining them. To characterize each pattern in terms of its property values (i.e., its defining ranking conditions), we base ourselves on the formal model of a Basic Truncation Typology (BTT) proposed in Alber & Arndt-Lappe (2022). This model offers an intensional classification of truncation patterns according to their size and anchoring properties by extracting the defining ranking conditions (the properties) of the typology – recall the size and anchoring classification in Table 1.[11] The model uses a set of five constraints. Two of them refer to size, requiring the truncatum to be mono- or disyllabic (M.1S, M.2S), and three refer to anchoring properties, requiring the truncatum to anchor to the left or right edge of the base (ANCHL, ANCHR), or to its stressed syllable (MAXSTRESS; see Alber & Arndt-Lappe for definitions and details of the analysis). The analysis yields the five properties in (10). The names of properties are given in the format 'Name of property.value x/value y';[12] constraint classes are indicated in uppercase letters; the variables .dom and .sub identify the dominant/subordinate constraint of a class. Definitions of properties are read as follows: for the property value Edge.L to hold, the constraint ANCHL must dominate the constraint ANCHR. For the opposite value, Edge.R, to hold, ANCHR must dominate ANCHL (Alber & Prince 2015, 2021).

11. Alber & Kokkelmans (2022) undertook a similar comparison between truncation patterns in French and German, ordered by perceived age, and their differences in terms of property values. The difference with respect to the present study is that their data was not sufficient to establish statistically significant differences in perceived age between all investigated patterns.
12. Property names are prefixed with 'p.' in order to distinguish them from constraint names.

(10) Properties in the formal model of truncation BTT (Alber & Arndt-Lappe 2022)

Property	**Definition**	**extensional trait**
p.Edge.L/R	AnchL < > AnchR	edge-orientation: left- vs. right anchored
p.Prom.stress/edge	MaxStress < > EDGE.dom	stress vs. edge anchored
p.Trunc.noT/T	EDGE.sub < > TMPL.dom	no truncation vs. some truncation
p.Length.long/short	PROM.sub < > TMPL.dom	double vs. single anchoring = long truncation vs. short truncation
p.Tmpl.di/mono	m.2s < > m.1s	template size: mono- vs. disyllabic
Constraint classes: TMPL = {m.1s, m.2s} = the class of constraints demanding mono- or disyllabic shape EDGE = {AnchL, AnchR} = the class of constraints demanding left- and right-anchoring PROM = {MaxStress, EDGE.dom} = the class containing the constraint demanding stress-anchoring and the dominant of the two EDGE constraints		

Different ranking conditions are responsible for generating different classes of truncation patterns characterized by certain extensional traits. The property p.Edge decides whether a pattern is left- or right-anchored; the ranking defining p.Prom decides whether the truncatum is anchored to the stressed syllable or to some edge; p.Trunc tells us whether truncation takes place at all;[13] the ranking defining the property p.Length distinguishes long patterns like the one exemplified by *Antonélla – An.to.né* from short, templatic patterns like *Francésca – Frán.ce;* the ranking between m.1s and m.2s decides the value of the property p.Tmpl, which divides patterns in mono- and disyllabic ones.

Table (11) shows how the truncation patterns tested in our questionnaire can be classified according to the property values of the formal model BTT. All four patterns are specified for the property value Trunc.T since all of them show some degree of truncation. The pattern exemplified by *Francésca-Cés.ca* (Pattern D) is ambiguous in its classification. It could be classified as the mirror image of the *Antonélla-An.to.né* pattern, i.e., as a pattern anchoring both to the stressed and the final syllable. In this case, if would be specified for the property values Length. long and Edge.R. It is not possible to determine whether, under this classification,

13. TruncT.noT = EDGE.sub > TMPL.dom. This definition of the value 'no Truncation' states that no truncation occurs if the subordinate EDGE constraint, i.e., both EDGE constraints, dominate the dominant, i.e., both TMPL constraints. Anchoring to both edges makes truncation impossible in a model like BTT where deletion leaving internal gaps (hence violating the contiguity of segments) is not contemplated.

the property p.Tmpl is specified as mono- or di., since this difference emerges only in base names with final stress (e.g., hypothetical *Matará*) where a decision has to be made to copy either only the stressed final syllable (*Rá*; Tmpl.mono) or the stressed final plus its preceding syllable (*Tará*; Tmpl.di). No base names with final stress are attested in our sample. As to the value of the property p.Prom, if we classify *Cés.ca* as a pattern anchoring from the stressed syllable to the right edge, this property is moot: the pattern can be defined neither as stress- nor as edge-anchored since it is both stress- *and* edge-anchored. The second possible classification of *Cés.ca* is that of a short pattern (Length.short) which is disyllabic (Tmpl.di) and anchors to the stressed syllable (Prom.stress). Edge orientation is to the right (Edge.R), under this classification, since otherwise the output should be preserving the syllable to the left of the stressed syllable, not that to the right. The third possible classification of *Francésca-Cés.ca*, where this pattern is seen as a right-anchoring pattern, is defined by the same set of property values as the previous one except for the value of p.Edge, which is now Edge.edge, rather than Edge.stress. As discussed in Section 1, in languages where stress is by default on the penultimate syllable, it is not possible to distinguish between the three classifications of this pattern. Pattern B does not exhibit ambiguities of this type. It is classified as a disyllabic (Tmpl.di), edge-anchoring (Prom.edge) and left-oriented (Edge.L) pattern. Pattern A can be classified similarly, except for the fact that it takes the property value Tmpl.mono, which is responsible for its monosyllabicity. Pattern C is classified here as fulfilling a disyllabic template at the word-level (hence Tmpl.di) and is otherwise defined by the same property values as Pattern A.[14]

(11) Diachronic change of truncation patterns by minimal changes of property values

	old → recent					
	Pattern D			Pattern B	Pattern A	Pattern C
	Cés.ca (stress-to-right edge)	Cés.ca (stress)	Cés.ca (right edge)	Frán.ce	Fra	Frán.c-i
p.Trunc	T	T	T	T	T	T
p.Length	long	short	short	short	short	short
p.Tmpl	/	di	di	di	mono	di
p.Edge	R	R	R	L	L	L
p.Prom	--	stress	edge	edge	edge	edge

The left-to-right order in which the four patterns are presented in (12) corresponds to the order of perceived age D > B > A, form oldest to most recent,

14. An alternative would be to interpret Pattern C represented by *Frán.c-i* as satisfying a monosyllabic template at the level of the truncation morpheme, which then is suffixed by *-i*. For discussion of the two possible analyses see Féry (1997), Wiese (2001), Alber (2007), Arndt-Lappe (2018).

established in the experiment. As discussed in the previous section, the place of Pattern C in this hierarchy is not completely clear, since besides age other factors such as the gender of the name bearer seem to be at stake when it comes to the selection of short names of this type. We therefore tentatively place Pattern C at the margin of the hierarchy, separated from Pattern A by a dotted line. This placement recognizes the fact that Pattern C must be listed among 'younger' patterns, but at the same time cannot be defined by age alone. Pattern D can receive three possible classifications, according to the formal model; they are also divided by dotted lines. To verify our hypothesis that diachronic change in patterns proceeds in minimal intensional steps we must check whether the path from D to B to A does involve the change of one property value at a time. There is indeed a possible analysis of this type. The interpretation of Pattern D as a right-anchoring pattern, when compared to Pattern B, involves one change in property values: from Edge.R to Edge.L (shaded). All other property values remain the same. The change from Pattern B to Pattern A involves a single change in value as well, from a disyllabic (Tmpl.di) to a monosyllabic (Tmpl.mono) pattern. Pattern C, as long as it is ordered among the younger patterns to the right of Pattern D, will also involve at most a single change of property value, either from Edge.right to Edge.left (D --> C) or from Tmpl.mono to Tmpl.di (A --> C). No change in property values occurs from Pattern B to Pattern C or vice versa.

For the analysis to implement minimalism, *Cés.ca* must be interpreted as right- and not as stress-anchoring. However, also the interpretations of *Cés.ca* as stress-to-right-anchoring or as stress-anchoring can be accommodated in a minimal step analysis – if these two classifications are assumed to be older. They can be analyzed as proceeding from a long to a short pattern (Length.long/short) and then from a stress-anchored to an edge-anchored pattern (Prom.stress/edge). The analysis predicts, however, that there must have been a point in time where the pattern generating *Francésca-Cés.ca* was interpreted as right-anchoring, leading, in its next diachronic step, to the left-anchoring pattern *Francésca-Frán.ce.* This prediction could be verified by checking older attestations of short names for the rare base names bearing antepenultimate stress. We should for instance find in older sources short names like *Ní.co* for base names such as *Doménico.*

In sum, the diachronic order that we found among patterns in our experiment yields an analysis where the change from one pattern to the next involves a minimal change in grammar. The formal typological model allowed us to define the notion of minimal change in grammar in a very concrete way: grammars differ minimally when they differ in a single property value.

5. Conclusions

Word formation processes such as name truncation have been considered outside of the domain of grammar proper by some analysts (Dressler & Merlini-Barbaresi 1994; Dressler 2000; Mattiello 2013; Ronneberger-Sibold 2015). Besides the fact that it is difficult to pinpoint the exact meaning and function of the output of truncation and that this process is an instance of non-concatenative morphology, one

of the arguments that has been put forward in favor of the extragrammaticality of name truncations is their presumed unpredictability.

It is true that in many languages there is more than one way to shorten a name. However, once we recognize patterns of name truncation, classifiable in terms of size and anchoring, it is possible to distinguish these patterns in terms of their meaning and function and in terms of their variation in time and space.

In this paper we focused on the coexistence of several truncation patterns in the same language as the result of language change. It seems that in some languages the turnover of patterns is rather fast, leading to the simultaneous presence of patterns perceived as older and patterns perceived as more recent. We studied the age of truncation patterns in the Romance and Germanic varieties of Northern Italy using a novel methodology, asking speakers which short name they consider appropriate for the persona in a dialogue belonging to a certain age group. The use of a crowdsourcing platform to carry out the experiment made it possible to reach many speakers spread over a large geographical space. The results obtained show with statistical significance that speakers do discriminate between older and more recent patterns. It has shown also for one pattern, the *i*-suffixed Pattern C represented by *Francésca – Frán.ci*, which was preferentially assigned to female name bearers, that sometimes there is more than age to a pattern and that these additional factors can be uncovered. This means that, at least for these languages and these patterns, the output of name truncation for a certain name is predictable. If a speaker uses a short name for their friend *Francésca*, the type of short name they use will depend to a large extent on the age of *Francésca.* If Francesca is rather young, it is probable that she will be called *Fra* or *Fran.ce*, or, since this is a female name, *Frán.c-i.* If she belongs to the older generation, she might be called *Cés.ca.*

We showed that Italian and German short names are predictable also in a second sense. If it is true that the domain of prosodic morphology language change proceeds in small steps (one typological property at a time) then the order of patterns according to age is not random. Not every change is possible, only those instantiating minimal changes are, hence not every pattern is predicted to coexist with every other pattern to the same degree of probability.

As is always the case, this study does not exhaust the interesting lines of research. We have shown that language change plays a role in the name truncation patterns of Italian and German as spoken in Northern Italy (see Alber & Kokkelmans 2022 for a similar conclusion for French). In other languages and other speech communities, different factors may play a role in favoring the coexistence of more than one name truncation pattern. Differences in meaning and function of truncation patterns remain an understudied area of research. Open questions also remain with respect to how truncation relates to similar phenomena like clipping and blends. If language change plays a role in the emergence of name truncation patterns, we would furthermore expect to find languages where the spread of innovative pattern can be traced in space. For Northern Italy we did find first indications that this is indeed the case, since innovative patterns seem to be attested more in the West, where cities such as Milan and Turin are located. With the tools of geolinguistic analysis it should be able to detect variation of this type also for the patterns

of short names. All these lines of research can of course be undertaken only under the assumption that there is structure to be detected – and not unpredictable chaos beyond the regularities of grammar.

References

#TrentoInCifre 2023 = #TrentoInCifre. *La popolazione di Trento al 31 dicembre 2023*. Ufficio Studi e Statistica, Comune di Trento. <https://www.comune.trento.it/Aree-tematiche/Statistiche-e-dati-elettorali/Statistiche/Pubblicazioni/Popolazione-al-31-dicembre/Popolazione-al-31-dicembre-2023>

Abuladze, Lia & Ludden, Andreas. 2013. The vocative in Georgian. In Sonnehauser, Barbara & Noel Aziz Hanna, Patrizia (eds.). *Vocative! Adressing between system and performance*, 25-42. Berlin/Boston: De Gruyter.

Alber, Birgit. 2007. Deutsche und italienische Kurzwörter im Vergleich. In Di Meola, Claudio Gaeta, Livio, Hornungm Antonie & Rega, Lorenza (eds.). *Perspektiven Zwei. Akten der 2. Tagung Deutsche Sprachwissenschaft in Italien* (Rom, 9.-11.2. 2006), 101-112. Roma: Istituto Italiano di Studi Germanici.

Alber, Birgit. 2010. An exploration of truncation in Italian. In Staroverov, Peter, Altshuler, Daniel, Braver, Aaron, Fasola, Carlos A. & Murray, Sarah (eds.). *Rutgers Working Papers in Linguistics*, 1-30. New Brunswick: LGSA.

Alber, Birgit & Arndt-Lappe, Sabine. 2012. Templatic and subtractive truncation. In Trommer, Jochen (ed.). *The Phonology and Morphology of Exponence - the State of the Art* (Oxford Studies in Theoretical Linguistics), 289-325. Oxford: Oxford University Press.

Alber, Birgit & Arndt-Lappe, Sabine. 2022. Anchoring in Truncation: a typological analysis. *Natural Language and Linguistic Theory* 41: 1-50. <https://doi.org/10.1007/s11049-021-09534-x>

Alber, Birgit & Arndt-Lappe, Sabine. 2023. Clipping and Truncation. In Ackema, Peter, Benjaballah, Sabrina, Bonet, Eulàlia & Fábregas, Antonio (eds.). *The Wiley Blackwell Companion to Morphology*, 1-28. Chichester: Wiley Blackwell. <https://doi.org/10.1002/9781119693604.morphcom017>

Alber, Birgit & Kokkelmans, Joachim. 2022. Typology and language change: The case of truncation. *Isogloss* 8(2)/13: 1-17. <https://doi.org/10.5565/rev/isogloss.124>

Alber, Birgit & Meneguzzo, Marta. 2016. Germanic and Romance onset clusters - how to account for microvariation. In Bidese, Ermenegildo, Cognola, Federica & Moroni, Manuela C. (eds.). *Theoretical Approaches to Linguistic Variation*, 25-51. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.234.02alb>

Alber, Birgit & Prince, Alan. 2015. *Outline of Property Theory*. Ms., University of Verona/ Rutgers University.

Alber, Birgit & Prince, Alan. 2021. *The Structure of OT Typologies. Chapter 1: Introduction to Property Theory*. <https://roa.rutgers.edu/content/article/files/1838_alber_2.pdf>

Arndt-Lappe, Sabine. 2018. Expanding the Lexicon by Truncation: Variability, Recoverability, and Productivity. In Arndt-Lappe, Sabine, Braun, Angelika, Moulin, Claudine & Winter-Froemel, Esme (eds.). *Expanding the Lexicon: Linguistic*

Innovation, Morphological Productivity, and Ludicity, 141-170. Berlin/New York: De Gruyter Mouton.

Apostolopoulou, Eirini. 2022. *Typological variation in language contact. A phonological analysis of Italiot Greek*. PhD diss. University of Verona/University of Tromsø – The Arctic University of Norway.

ASTAT 2018 = Autonome Provinz Bozen-Südtirol, Landesinstitut für Statistik / Istituto provinciale di statistica. 2018. *Vornamen in Südtirol / Nomi propri in provincia di Bolzano*. Bozen-Bolzano. <https://astat.provinz.bz.it/>

Boschiroli, Laura. 2017. *Bestimmung des Alters italienischer Trunkierungsmuster*. Term paper, University of Verona.

Brattö, Olaf. 1953. *Studi di antroponimia fiorentina. Il Libro di Montaperti (An. MCCLX)*. Göteborg: Elander.

D'Alessandro, Roberta & van Oostendorp, Marc. 2016. When imperfections are perfect: Prosody, phi-features and deixis in Central and Southern Italian vocatives. In Carrilho, Ernestina, Fiéis, Alexandra, Lobo, Maria & Pereira, Sandra (eds.). *Romance Languages and Linguistic Theory 10: Selected Papers from Going Romance 28, Lisbon*, 61-82. Amsterdam: Benjamins.

Dressler, Wolfgang Ulrich. 2000. Extragrammatical vs. Marginal Morphology. In Doleschal, Ursula & Thornton, Anna Maria (eds.). *Extragrammatical and Marginal Morphology* (LINCOM Studies in Theoretical Linguistics 12), 61-82. München: LINCOM Europa.

Dressler, Wolfgang Ulrich & Merlini-Barbaresi, Lavinia. 1994. *Morphopragmatics: Diminutives and Intensifiers in Italian, German, and Other Languages*. Berlin/New York: Mouton De Gruyter.

Féry, Caroline. 1997. Uni und Studis: Die besten Wörter des Deutschen. *Linguistische Berichte* 172: 461-489.

Germain, Jean. 2014. La loi des deux syllabes: une réponse structurelle à la fixation des noms de famille en Wallonie. In Gilles, Peter, Kollmann, Cristian & Muller, Claire (eds.). *Familiennamen zwischen Maas und Rhein* (Études luxembourgeoises/ Luxemburg-Studien 6), 141-155. Frankfurt/Main et al.: Peter Lang.

Kenstowicz, Michael. 2019. The analysis of truncated vocatives in Taviano (Salentino) Italian. *Catalan Journal of Linguistics* 18: 131-159.
<https://doi.org/10.5565/rev/catjl.257>

Mattiello, Elisa. 2013. *Extra-grammatical Morphology in English. Abbreviations, Blends, Reduplicatives, and Related Phenomena*. Berlin: De Gruyter.

Merlini Barbaresi, Lavinia & Dressler, Wolfgang Ulrich. 2020. Pragmatic Explanations in Morphology. In Pirelli, Vito, Plag, Ingo & Dressler, Wolfgang Ulrich (eds.). *Word Knowledge and Word Usage*, 406-451. Berlin/Boston: De Gruyter Mouton.
<https://doi.org/10.1515/9783110440577-011>

Prince, Alan & Smolensky, Paul. 2004 [1993]. *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwell.

Rabanus, Stefan, Kruijt, Anne, Alber, Birgit, Bidese, Ermenegildo, Gaeta, Livio & Raimondi, Gianmario. 2024. *AlpiLinK Corpus 1.1.1*. In collaboration with Benedetto Mas, Paolo, Bertollo, Sabrina, Bonelli, Angelica, Casalicchio, Jan, Cioffi, Raffaele, Cordin, Patrizia, Dal Negro, Silvia, Glück, Alexander, Kokkelmans, Joachim, Murelli, Adriano, Padovan, Andrea, Pons, Aline, Rivoira, Matteo,

Tagliani, Marta, Saracco, Caterina, Tomaselli, Alessandra, Videsott, Ruth, Vietti, Alessandro & Vogt, Barbara.
<https://doi.org/10.5281/zenodo.11083948>

Raus, Sabrina. 2024. *Italian and German name truncation in Trentino-Alto Adige*. MA-thesis, Free University of Bozen-Bolzano.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rohlfs, Gerhard. 1949. *Historische Grammatik der Italienischen Sprache und ihrer Mundarten. Band 1: Lautlehre*. Bern: Francke Verlag.

Ronneberger-Sibold, Elke. 2015. Word Creation. In Müller, Peter O., Ohnheiser, Ingeborg, Olsen, Susan, & Rainer, Franz (eds.). *Word-Formation – An International Handbook of the Languages of Europe 2*: 485-499. Berlin: De Gruyter.
<https://doi.org/10.1515/9783110246254-028>

Scherillo, Michele. 1896. *Alcuni capitoli della biografia di Dante*. Torino: Ermanno Loescher.

Schneider, Klaus Peter. 2003. *Diminutives in English*. Tübingen: Niemeyer.

Selkirk, Elisabeth. 2011. The syntax-phonology interface. In Goldsmith, John A., Riggle, Jason & Yu, Alan C. L. (eds.). *The Handbook of Phonological Theory*, 435-484. Oxford: Blackwell.

Venables, William N. & Ripley, Brian D. 2002. *Modern Applied Statistics with S*. Fourth edition. New York: Springer.

Wiese, Richard. 2001. Regular Morphology vs. Prosodic Morphology? The Case of Truncations in German. *Journal of Germanic Linguistics* 13(2): 131-177.

Appendix

Table 1. z-values and p-values of comparisons for the individual base name models presented in fig. (5) and (6): NameAgeSetting

Theresa				
young vs. mid		B	C	D
	A vs.	z = -0.37, p = 0.71	z = 0.35, p = 0.72	z = 0.49, p = 0.62
	B vs.		z = 0.82, p = 0.41	z = 0.98, p = 0.33
	C vs.			z = 0.15, p = 0.88
young vs. old		B	C	D
	A vs.	z = -1.05, p = 0.29	z = -0.88, p = 0.38	z = 3.62, p < 0.01
	B vs.		z = 0.24, p = 0.81	z = 4.60, p < 0.01
	C vs.			z = 4.96, p < 0.01
Isabella				
young vs. mid		B	C	D
	A vs.			
	B vs.		z = -0.21, p = 0.83	z = -0.66, p = 0.51
	C vs.			z = -0.51, p = 0.61
young vs. old		B	C	D
	A vs.			
	B vs.		z = -2.39, p = 0.02	z = 0.90, p = 0.37
	C vs.			z = 2.63, p = 0.01
Matthias				
young vs. mid		B	C	D
	A vs.			z = 4.57, p < 0.01
	B vs.			
	C vs.			
young vs. old		B	C	D
	A vs.			z = 5.88, p < 0.01
	B vs.			
	C vs.			

Giovanni				
young vs. mid		B	C	D
	A vs.	z = 3.23, p < 0.01	z = -0.65, p = 0.51	z = 6.57, p < 0.01
	B vs.		z = -2.94, p < 0.01	z = 4.28, p < 0.01
	C vs.			z = 6.09, p < 0.01
young vs. old		B	C	D
	A vs.	z = 5.48, p < 0.01	z = 1.11, p = 0.27	z = 10.4, p < 0.01
	B vs.		z = -3.1, p < 0.01	z = 6.39, p < 0.01
	C vs.			z = 8.15, p < 0.01
Francesca				
young vs. mid		B	C	D
	A vs.	z = 3.07, p < 0.01	z = -1.54, p = 0.12	z = -1.73, p = 0.08
	B vs.		z = -4.25, p < 0.01	z = -4.23, p < 0.01
	C vs.			z = -0.60, p = 0.55
young vs. old		B	C	D
	A vs.	z = 5, p < 0.01	z = 0.58, p = 0.56	z = 6.63, p < 0.01
	B vs.		z = -5.29, p < 0.01	z = -0.57, p = 0.57
	C vs.			z = 8.71, p < 0.01
Matteo				
young vs. mid		B	C	D
	A vs.			
	B vs.		z = -2.09, p = 0.04	z = 2.44, p = 0.01
	C vs.			z = 4.02, p < 0.01
young vs. old		B	C	D
	A vs.			
	B vs.		z = -0.79, p = 0.43	z = 7.86, p < 0.01
	C vs.			z = 6.78, p < 0.01

Table 2. z-values and p-values of comparisons for the individual base name models presented in fig. (5) and (6): OwnAge

Theresa				
OwnAge		B	C	D
	A vs.	z = -2.68, p = 0.01	z = -0.65, p = 0.52	z = 1.19, p = 0.24
	B vs.		z = 2.46, p = 0.01	z = 4.43, p < 0.01
	C vs.			z = 2.30, p = 0.02
Isabella				
OwnAge		B	C	D
	A vs.			
	B vs.		z = 0.93, p = 0.35	z = 1.57, p = 0.12
	C vs.			z = 0.85, p = 0.40
Matthias				
OwnAge		B	C	D
	A vs.			z = 1.01, p = 0.31
	B vs.			
	C vs.			
Giovanni				
OwnAge		B	C	D
	A vs.	z = -4.61, p < 0.01	z = -2.15, p = 0.03	z = -0.56, p = 0.58
	B vs.		z = 1.47, p = 0.14	z = 3.03, p < 0.01
	C vs.			z = 1.29, p = 0.20
Francesca				
OwnAge		B	C	D
	A vs.	z = 1.17, p = 0.24	z = 0.75, p = 0.45	z = 5.02, p < 0.01
	B vs.		z = -0.77, p = 0.44	z = 2.76, p = 0.01
	C vs.			z = 5.68, p < 0.01
Matteo				
OwnAge		B	C	D
	A vs.			
	B vs.		z = 7.08, p < 0.01	z = 6.38, p < 0.01
	C vs.			z = -2.70, p = 0.01

Table 3. z-values and p-values of comparisons for the BCD-model presented in fig. (7)

young vs. mid		C	D
	B vs.	z = -2.91, p < 0.01	z = 1.72, p = 0.09
	C vs.		z = 4.77, p < 0.01
young vs. old		C	D
	B vs.	z = -5.44, p < 0.01	z = 9.19, p < 0.01
	C vs.		z = 15.04, p < 0.01
OwnAge		C	D
	B vs.	z = 4.14, p < 0.01	z = 8.89, p < 0.01
	C vs.		z = 4.35, p < 0.01
NameGender: F vs. M		C	D
	B vs.	z = -23.16, p < 0.01	z = -8.26, p < 0.01
	C vs.		z = 19.30, p < 0.01
Germanic vs. Romance		C	D
	B vs.	z = 15.27, p < 0.01	z = 8.99, p < 0.01
	C vs.		z = -6.78, p < 0.01
F vs. M * Germanic vs. Romance		C	D
	B vs.	z = -23.16, p < 0.01	z = -8.26, p < 0.01
	C vs.		z = 19.30, p < 0.01

Table 4. z-values and p-values of comparisons for the longitude and latitude model presented in fig. (8)

Longitude			
	B	C	D
A vs.	z = 4.84, p < 0.01	z = 7.11, p < 0.01	z = 8.09, p < 0.01
B vs.		z = 1.50, p = 0.13	z = 2.62, p < 0.01
C vs.			z = 1.47, p = 0.14
Latitude			
	B	C	D
A vs.	z = -0.57, p = 0.57	z = -0.80, p = 0.42	z = -13.33, p < 0.01
B vs.		z = -0.23, p = 0.82	z = -10.81, p < 0.01
C vs.			z = -16.00, p < 0.01