PFA RESEARCH PAPER 001

PHONETIC FLUIDITY AND THE SPOKEN DOMAIN NAME

Investigating whether phonetic and structural characteristics may influence the processing of unfamiliar domain names

PFA Research Programme
Research Paper 001
Initial Research Publication

Status: Research Record

domainAlot.com PREMIUM DOMAINS YOU CAN TRUST

ABSTRACT

Domain names have traditionally been evaluated primarily as written identifiers.

Their length, memorability, semantic meaning, visual appearance, extension, market relevance and other characteristics have all been considered in different approaches to domain assessment.

The increasing use of voice interfaces, conversational systems and speech-based interaction introduces another environment in which a domain name may be encountered: the spoken environment.

This paper investigates a foundational question underlying the Phonetic Fluidity Assessment (PFA):

> **To what extent does the phonetic and structural character of a domain name affect the ease with which the name can be spoken, heard, recognised, remembered or reproduced?**

The paper reviews relevant research concerning phonological structure, phonotactic probability, pronunciation fluency, unfamiliar names, novel brand names, auditory-to-written conversion, sound symbolism and automatic speech recognition.

The evidence provides reasonable support for the broader proposition that characteristics of linguistic form can influence aspects of human processing and that unfamiliar or rare names can present challenges for speech-recognition systems.

However, the evidence does not establish that the specific characteristics, weighting, scoring system or Tier classifications used by PFA predict those outcomes for domain names.

The findings therefore support **further investigation of the underlying phenomenon**, but do not constitute validation of the PFA Framework itself.

This distinction is fundamental.

1. INTRODUCTION

A domain name is normally encountered as written text.

A person sees it on a website, advertisement, search result, document, email, presentation or other visual medium. Traditional domain evaluation has therefore naturally focused on characteristics that can be observed in written form.

The communication environment surrounding domain names is changing.

Voice assistants, speech interfaces, conversational AI and other systems increasingly allow people to interact with digital services through spoken language. At the same time, users can encounter brand and domain names through conversations, podcasts, video, meetings, telephone calls, advertising and other primarily auditory environments.

A spoken domain therefore creates a different communication problem from a visually encountered domain.

The listener may need to:

1. Hear the name accurately.
2. Segment its sounds.
3. Reproduce the pronunciation.
4. Remember the name.
5. Infer or reproduce its written representation.
6. Communicate the name to another person or system.
7. Use the resulting representation to locate the intended digital destination.

These processes involve language, perception, memory, pronunciation and, increasingly, speech-recognition technology.

The **PFA Research Programme** was established partly from the proposition that these characteristics may deserve systematic examination.

The purpose of this paper is not to assume that proposition is correct.

It is to investigate whether existing evidence provides a reasonable foundation for asking the question.

2. RESEARCH QUESTION

The primary research question is:

> **To what extent does the phonetic and structural character of a domain name affect the ease with which the name can be spoken, heard, recognised, remembered or reproduced?**

Five supporting questions are examined:

2.1 PRONUNCIATION

Do particular phonetic structures make unfamiliar names more difficult to pronounce or repeat?

2.2 PERCEPTION AND RECOGNITION

Can phonological characteristics influence how readily unfamiliar spoken forms are processed or recognised?

2.3 MEMORY

Can pronunciation, phonological structure or related linguistic characteristics influence retention or recall?

2.4 SPEECH RECOGNITION

Do rare, unfamiliar or structurally unusual names create measurable challenges for automatic speech-recognition systems?

2.5 GENERALISABILITY

Can findings concerning phonetic or linguistic processing reasonably be generalised across languages, accents, dialects and communication environments?

3. SCOPE OF THIS PAPER

This paper examines the **underlying phenomenon** that motivated the development of PFA.

It does not attempt to establish that:

- PFA predicts domain value.
- A particular PFA Score predicts commercial success.
- A particular PFA Tier represents objectively superior economic value.
- A domain with a lower PFA Score will necessarily be harder for a voice assistant to recognise
- PFA improves traffic, conversion or revenue.
- PFA replaces established approaches to domain valuation.

Most importantly, this paper does not attempt to validate the specific PFA scoring algorithm.

Those are separate research questions.

The distinction is intentional.

4. WORKING DEFINITION OF PHONETIC FLUIDITY

"Phonetic Fluidity" is used within the PFA Framework as a working term describing characteristics that may influence how naturally or easily a name can move through spoken communication.

The term should not be interpreted as an established scientific construct merely because it is used by PFA.

Existing research uses a number of related but distinct concepts, including:

- Phonotactic probability.
- Pronunciation difficulty.
- Processing fluency.
- Phonological structure.
- Sound-to-spelling consistency.
- Lexical familiarity.
- Phonological neighbourhood.
- Speech-recognition accuracy.
- Sound symbolism.

These concepts overlap in some respects but are not interchangeable.

Consequently, one of the questions raised by this research is whether "Phonetic Fluidity" should ultimately be treated as a single construct or as a composite description covering several measurable properties.

At this stage, PFA treats it as a **working construct requiring further investigation**.

5. EVIDENCE AREA I: PHONOLOGICAL STRUCTURE AND PROCESSING

A substantial body of psycholinguistic research has examined how the probability of particular sound sequences within a language can influence language processing.

Phonotactic probability concerns how frequently particular phonological sequences occur within a language. Research has associated phonological probability with differences in the processing of spoken and novel linguistic forms.

This provides an important theoretical basis for investigating whether phonological structure matters when people encounter unfamiliar names.

However, phonotactic probability should not be equated with PFA's concept of Phonetic Fluidity.

A sequence may be statistically common without necessarily being easy for every speaker in every context.

Similarly, a low-probability sequence is not necessarily difficult to pronounce or recognise.

The evidence therefore supports the more limited proposition that:

Phonological structure can influence aspects of linguistic processing.

It does not establish:

A particular PFA structural score measures that influence accurately.

That distinction must remain explicit.

6. EVIDENCE AREA II: UNFAMILIAR NAMES

Names provide an especially relevant comparison because they are frequently encountered as linguistic forms that may have little or no semantic familiarity to the listener.

Research comparing surnames with ordinary nouns has found differences in familiarity and pronounceability, with pronounceability also relating to the time required to produce a name.

This suggests that unfamiliar names can impose processing characteristics that differ from those associated with familiar vocabulary.

That observation is relevant to domain names because many domain names are:

- Invented.
- Brand-like.
- Unfamiliar.
- Non-dictionary forms.
- Combinations of linguistic elements.
- New identifiers encountered without prior exposure.

A listener encountering an unfamiliar domain therefore cannot necessarily rely upon the same lexical familiarity available when processing a common word.

However, domains differ substantially from human names, and the existing evidence should not be transferred without qualification.

The appropriate conclusion is that **unfamiliar naming forms provide a legitimate research analogue**, not that research on surnames directly validates PFA.

7. EVIDENCE AREA III: BRAND NAMES AND NOVEL WORDS

Research into brand names provides another relevant body of evidence.

Brand names frequently include invented or non-dictionary forms, making them closer to domain names than ordinary vocabulary in some respects.

Research on fictitious brand names has examined the relationship between pronunciation fluency, memory and consumer judgments. One study reported effects of linguistic fluency on liking and memory of non-word brand names, while also finding that fluency could influence judgments concerning innovative products.

Other research has examined sound symbolism in brand names and found that particular phonetic characteristics can influence associations and perceptions.

A systematic review of sound symbolism in brand names concluded that particular phonetic elements can convey associations related to dimensions such as evaluation and potency.

These findings are relevant because they demonstrate that the sounds composing a novel name need not be psychologically neutral.

However, s**ound symbolism and Phonetic Fluidity are different concepts**.

Sound symbolism concerns associations between sounds and meanings or attributes.

Phonetic Fluidity, as used by PFA, concerns the potential ease with which a spoken name may be produced, perceived or reproduced.

The two may interact, but evidence for one should not automatically be treated as evidence for the other.

8. EVIDENCE AREA IV: AUDITORY NAMES AND SOUND-TO-SPELLING CONVERSION

One of the most directly relevant areas of research concerns what happens when a person hears an unfamiliar name and must reconstruct its written representation.

This problem is particularly relevant to domains.

A person may hear a domain name without seeing it.

They may then attempt to:

- Remember the name.
- Type it into a browser.
- Search for it.
- Communicate it to another person.
- Reproduce its spelling.

Research published in the *Journal of Consumer Psychology* examined how people process auditorily presented brand names and how spelling-related characteristics affect memory.

The researchers described both lexical and sublexical routes involved in converting sound into written representation and found that spelling-related characteristics could influence subsequent memory for brand names.

The research is particularly relevant to domain names because a spoken domain often ultimately needs to become a written domain string.

This creates an important distinction.

A name can be:

Easy to pronounce but difficult to spell.

Moreover, it can also be:

Easy to spell once seen but difficult to reconstruct from sound.

These are not necessarily the same problem.

This observation is highly relevant to the PFA concept sometimes described informally as a communication or “spelling” burden.

The research supports the existence of a **sound-to-spelling problem**, but it does not justify assigning a fixed financial cost to that problem.

Any such financial interpretation would require separate empirical evidence.

9. EVIDENCE AREA V: SPEECH RECOGNITION

The development of modern speech-recognition systems introduces another dimension.

Automatic speech-recognition systems do not process every word or name with equal reliability.

Research has specifically identified rare names, organisations and specialised terminology as challenging recognition targets.

For example, research involving contextual biasing for Whisper-based speech recognition has addressed the difficulty of recognising rare or unseen named entities and demonstrated that additional contextual information can improve recognition performance.

This provides meaningful evidence that:

> **Rare or unfamiliar linguistic entities can present recognition challenges to speech-recognition systems.**

This is directly relevant to the broader PFA Research Programme as well as the PFA Research question.

A domain name is often a rare or previously unseen string.

However, the existing ASR evidence also reveals an important qualification.

Recognition systems can use contextual information to improve recognition of difficult entities.

Therefore, recognition difficulty is not necessarily an intrinsic and permanent property of a name.

It may depend upon:

- The speech model.
- Language.
- Pronunciation.
- Context.
- Available vocabulary.
- Contextual biasing.
- Acoustic conditions.
- Speaker characteristics.
- System configuration.
- Prior exposure.

This means PFA should not describe a domain as simply “AI-readable” or “AI-unreadable” without specifying the system and conditions involved.

The evidence supports a **conditional problem**, not a universal binary property.

10. EVIDENCE AREA VI: PRONUNCIATION FLUENCY AND HUMAN RESPONSE

Research also suggests that ease of pronunciation can influence human judgments and perceptions.

Studies involving unfamiliar brand names have reported relationships between name fluency and consumer perceptions.

More recent research has continued to investigate difficult-to-pronounce brand names and their effects on consumer responses.

This supports the broader concept that pronunciation characteristics can matter beyond the act of pronunciation itself.

However, the effects are not necessarily universal or linear.

Some research indicates that phonological fluency can affect different outcomes in different ways, including memory, preference and evaluation.

This reinforces an important methodological point:

> **"Processing difficulty" should not automatically be treated as one universal outcome.**

A name may be:

- Difficult to pronounce without being difficult to remember.
- Easy to pronounce but difficult to spell.
- Unfamiliar but highly distinctive.
- Unusual yet memorable precisely because it is unusual.

These possibilities prevent a simplistic interpretation of Phonetic Friction.

11. EVIDENCE AREA VII: LANGUAGE AND CROSS-LINGUISTIC VARIATION

Any theory concerning phonetic structure must account for linguistic variation.

A sequence that is natural in one language may be unusual in another.

A consonant combination that is easy for one speaker may be difficult for another.

Pronunciation may vary according to:

- Language.
- Dialect.
- Accent.
- Phonological inventory.
- Orthographic conventions.
- Speaker experience.
- Familiarity with the name.

Research into phonetic symbolism has found effects across multiple languages, but cross-linguistic evidence does not establish that every phonetic effect is universal.

Other research has specifically demonstrated that phonological processing can depend upon language-specific characteristics.

This is a significant limitation for PFA.

A scoring system designed primarily around English-language phonological assumptions cannot automatically be treated as a universal model of spoken-domain performance.

Cross-linguistic validation therefore represents an important future research requirement.

12. WHAT THE EVIDENCE SUPPORTS

Taken together, the literature provides reasonable support for several propositions relevant to the PFA Research Programme.

PROPOSITION 1

Phonological characteristics can influence aspects of human language processing.

The evidence supports this proposition.

PROPOSITION 2

Pronunciation and linguistic fluency can influence how unfamiliar names are processed or evaluated.

The evidence provides meaningful support, although effects vary by outcome and context.

PROPOSITION 3

Unfamiliar names can create sound-to-spelling challenges.

The evidence supports this, particularly where multiple written representations are compatible with the heard form.

PROPOSITION 4

Rare or unfamiliar entities can create challenges for automatic speech recognition.

The evidence supports this, while also demonstrating that context and system design can mitigate those difficulties.

PROPOSITION 5

The characteristics of a spoken name may therefore deserve investigation when evaluating digital identifiers intended to operate partly through spoken communication.

The evidence provides a reasonable basis for investigating this proposition.

13. WHAT THE EVIDENCE DOES NOT ESTABLISH

The same evidence does not establish several stronger propositions.

It does not establish that:

PFA Scores predict speech-recognition accuracy.

No reviewed study directly demonstrates this.

It does not establish that:

PFA Tiers correspond to objectively different levels of spoken performance.

This remains unvalidated.

It does not establish that:

The specific PFA weighting of vowels, consonants, clusters, friction characters, plosives, length or other characteristics is empirically optimal.

These are currently methodological design choices.

It does not establish that:

A PFA Score predicts or otherwise guarantees commercial success.

That would require a separate body of empirical research.

It does not establish that:

A phonetic characteristic has a fixed monetary cost.

No such financial conversion follows from the linguistic evidence reviewed here.

It does not establish that:

A lower PFA Score means that every human or AI listener will experience greater difficulty.

Processing depends upon context, language, familiarity and system characteristics.

These boundaries are not weaknesses in the research.

They define the questions that remain open for further investigation.

14. THE CONCEPT OF PHONETIC FLUIDITY REQUIRES FURTHER DEFINITION

One of the most important findings from this review is that the term “Phonetic Fluidity” appears to combine several potentially separable phenomena.

These may include:

- Pronunciation Ease.
- Phonotactic probability.
- Sound-to-spelling consistency.
- Phonological complexity.
- Unfamiliarity.
- Processing fluency.
- Speech-recognition performance.
- Memorability.

These should not automatically be treated as one variable.

A future empirical programme may determine that some of these characteristics are strongly related, while others are relatively independent.

Consequently, the PFA Research Programme should not assume that a single numerical score necessarily represents one naturally occurring psychological or linguistic property.

An alternative possibility is that PFA is a **composite assessment framework** that combines several measurable characteristics associated with spoken-name performance.

This question requires further investigation.

15. IMPLICATIONS FOR THE PFA FRAMEWORK

The evidence reviewed in this paper provides a reasonable basis for continuing research into phonetic characteristics of domain names. However, at this stage, it does not yet justify treating the current PFA Algorithm as validated, which creates several priorities for future work.

15.1 CONSTRUCT DEFINITION

The PFA Research Programme should establish more precisely what PFA means by "Phonetic Fluidity" and how the concept relates to established linguistic constructs.

15.2 FEATURE VALIDATION

Individual characteristics currently considered by PFA should be tested independently.

15.3 WEIGHTING VALIDATION

The relative importance assigned to different characteristics should eventually be empirically examined rather than treated as inherently correct.

15.4 HUMAN TESTING

Controlled experiments could examine whether independently measured PFA characteristics correspond with pronunciation, recognition, recall or reproduction performance.

15.5 SPEECH-RECOGNITION TESTING

Domains could be tested across speech-recognition systems under controlled conditions.

15.6 CROSS-LINGUISTIC TESTING

The framework should eventually be examined across multiple languages, accents and pronunciation environments.

15.7 COMPARATIVE MODELS

PFA should ultimately be compared against simpler or alternative measures to determine whether its additional complexity provides measurable benefit.

These represent research opportunities rather than conclusions.

16. AN IMPORTANT ALTERNATIVE EXPLANATION

One of the most important questions raised by the evidence is whether **familiarity** may account for effects that might otherwise be attributed to phonetic structure.

A familiar name can be easier to process because the listener already knows it.

An unfamiliar name may appear difficult because it has never been encountered before.

That means an apparent relationship between phonetic structure and processing may sometimes be partly explained by:

- Lexical familiarity.
- Frequency.
- Prior exposure.
- Semantic associations.
- Spelling familiarity.
- Language familiarity.

Future PFA research must therefore control for these variables wherever possible.

Otherwise, it may incorrectly attribute a familiarity effect to phonetic structure.

This is a central methodological risk.

17. ANOTHER IMPORTANT DISTINCTION: HUMAN PROCESSING AND MACHINE PROCESSING

The evidence also demonstrates that human language processing and automatic speech recognition should not be treated as identical phenomena.

Humans and machines may encounter the same spoken name but process it through fundamentally different mechanisms.

Human listeners can use:

- Context.
- Semantic knowledge.
- Social cues.
- Conversational expectations.
- World knowledge.
- Speaker identity.
- Prior exposure.

Speech-recognition systems may use:

- Acoustic models.
- Language models.
- Contextual biasing.
- Vocabulary information.
- Decoding strategies.
- Learned representations.

Therefore:

> **Evidence that a name is difficult for humans does not automatically establish that it is difficult for an ASR system.**

Likewise:

> **Evidence that an ASR system struggles with a rare entity does not automatically establish that human listeners will struggle with it.**

The PFA Research Programme should keep these two environments conceptually distinct.

18. THE "SPELLING TAX" QUESTION

The research reviewed here provides a useful basis for revisiting the concept sometimes described within PFA as a "Spelling Tax."

A spoken name can create additional effort when the listener must infer its written representation, particularly when multiple spellings are plausible.

Research on auditorily presented brand names supports the existence of this sound-to-spelling challenge.

However, the evidence does not establish a universal numerical cost.

It therefore would be inappropriate to claim that a particular spelling characteristic causes a fixed amount of lost traffic, revenue or commercial value.

The more defensible interpretation is:

> **A spoken name may impose additional communication effort when its auditory representation does not provide a sufficiently clear route to its intended written representation.**

This is a researchable proposition.

It is not currently a monetary measurement.

19. LIMITATIONS OF THIS PAPER

This paper has several important limitations.

First, the literature reviewed concerns related linguistic, psycholinguistic, branding and speech-recognition phenomena rather than domain names specifically.

Second, the studies examine different dependent variables. Pronunciation, recognition, memory, preference and speech-recognition accuracy are not equivalent outcomes.

Third, the evidence comes from different languages, populations, tasks and experimental conditions.

Fourth, many studies examine controlled experimental stimuli rather than naturally occurring domain names.

Fifth, the research reviewed here does not directly test the current PFA Scoring Algorithm.

Sixth, the current PFA concept of “Phonetic Fluidity” is a working construct rather than an established scientific measurement.

These limitations substantially constrain the conclusions that can currently be drawn.

20. CONCLUSION

The evidence reviewed in this paper supports an important but deliberately limited conclusion.

There is credible research demonstrating that phonological structure, phonotactic characteristics, pronunciation fluency, familiarity and sound-to-spelling relationships can influence aspects of human processing.

There is also evidence that unfamiliar and rare linguistic entities can present challenges for automatic speech-recognition systems, while contextual information can substantially alter recognition performance.

These findings provide a legitimate basis for investigating whether the spoken characteristics of domain names deserve consideration in a changing digital communication environment.

However, the evidence does **not** establish the PFA Framework itself.

It does not demonstrate that PFA's specific characteristics, scoring weights or Tier classifications predict pronunciation, human recognition, memory, sound-to-spelling accuracy or speech-recognition performance.

Nor does it establish any direct relationship between a PFA Score and economic value or commercial success.

The appropriate conclusion at this stage is therefore:

> **The underlying phenomenon motivating PFA is sufficiently supported by existing research to justify systematic investigation, but the PFA Framework itself remains an unvalidated research hypothesis requiring empirical testing.**

This is the boundary between what the evidence currently permits us to say and what remains to be demonstrated.

That boundary should remain explicit.

21. RESEARCH QUESTIONS ARISING FROM THIS PAPER

The review produces several questions for future investigation.

QUESTION 1

Can the concept of "Phonetic Fluidity" be operationalised using established linguistic and psycholinguistic measures?

QUESTION 2

Which measurable characteristics of a domain name most strongly influence human pronunciation and recognition?

QUESTION 3

Does PFA predict human performance better than simpler measures such as length, phonotactic probability or sound-to-spelling consistency?

QUESTION 4

Does PFA predict speech-recognition performance?

QUESTION 5

Do different speech-recognition systems produce different results for the same domain names?

QUESTION 6

How much does contextual information reduce recognition difficulty?

QUESTION 7

How do language, accent and dialect affect the relationship between domain structure and spoken performance?

QUESTION 8

Are PFA's current scoring weights empirically defensible?

QUESTION 9

Do the current PFA Tier boundaries correspond to measurable differences in performance?

QUESTION 10

Does PFA provide information beyond what can be obtained from established linguistic measures?

These questions indicate a potential research programme beyond Paper 001.

22. FINAL RESEARCH POSITION

PFA began with an observation:

> **If the role of the domain has changed, the characteristics by which we evaluate a domain may also need to evolve.**

The evidence examined here does not prove that this proposition is correct in every respect.

It does, however, suggest that the underlying question is legitimate.

Human language processing is influenced by characteristics of linguistic form.

Unfamiliar names can create distinctive processing challenges.

Auditory information may need to be converted into written representations.

Speech-recognition systems can struggle with rare and unseen entities.

These phenomena provide reasonable grounds for asking whether domain names should be examined partly through the characteristics of their spoken forms.

The next question is harder.

It is whether those characteristics can be measured reliably.

And harder still:

> **Can a structured measure such as PFA demonstrate that it measures something meaningful, consistently and predictively?**

That question remains open.

It should remain open until the evidence is able to provide sufficient answers to it.

REFERENCES

Erz, A., & Christensen, B. (2013).
Fluency of Brand Names: Effects of Ease-of-Pronunciation on Non-Word Memory and Product Judgments. European Advances in Consumer Research, 10, 196-197.

Kuehnl, C., & Mantau, A. (2013).
Same sound, same preference? Investigating sound symbolism effects in international brand names. International Journal of Research in Marketing, 30(4), 417-420.

Luna, D., Carnevale, M., & Lerman, D. (2013).
Does brand spelling influence memory? The case of auditorily presented brand names. Journal of Consumer Psychology, 23(1), 36-48. DOI: 10.1016/j.jcps.2012.02.003.

Motoki, K., Park, J., Pathak, A., & Spence, C. (2022).
The connotative meanings of sound symbolism in brand names: A conceptual framework. Journal of Business Research, 150, 365-373. DOI: 10.1016/j.jbusres.2022.06.013.

Li, Y., Li, Y., Zhang, M., Su, C., Yu, J., Piao, M., Qiao, X., Ma, M., Zhao, Y., & Yang, H. (2024).
CB-Whisper: Contextual Biasing Whisper Using Open-Vocabulary Keyword-Spotting.
Proceedings of LREC-COLING 2024, 2941-2946.

Fenko, A., Lotterman, H., & Galetzka, M. (2016).
What's in a name? The effects of sound symbolism and package shape on consumer responses to food products. Food Quality and Preference, 51, 100-108.

[Additional phonotactic and psycholinguistic literature]
Further primary literature concerning phonotactic probability, pronunciation, lexical processing, unfamiliar names and related constructs should be incorporated into the formal evidence library as the PFA Research Programme continues its literature review.

RESEARCH STATUS

This paper represents the initial research position of the PFA Research Programme on the relationship between phonetic characteristics and spoken domain names.

It should not be interpreted as validation of the PFA Framework.

Future evidence may strengthen, qualify, contradict or replace conclusions presented here.

If subsequent research materially changes the position established in this paper, that development should be documented within the PFA Research Programme.

Research should remain capable of changing its answer.