



Language Properties of Phone Scammers: Cyberdefense at the Level of the Human

Aug. 4, June 2016

Judith Tabron, Ph.D.
Hofstra University



HOFSTRA
UNIVERSITY

Overview:

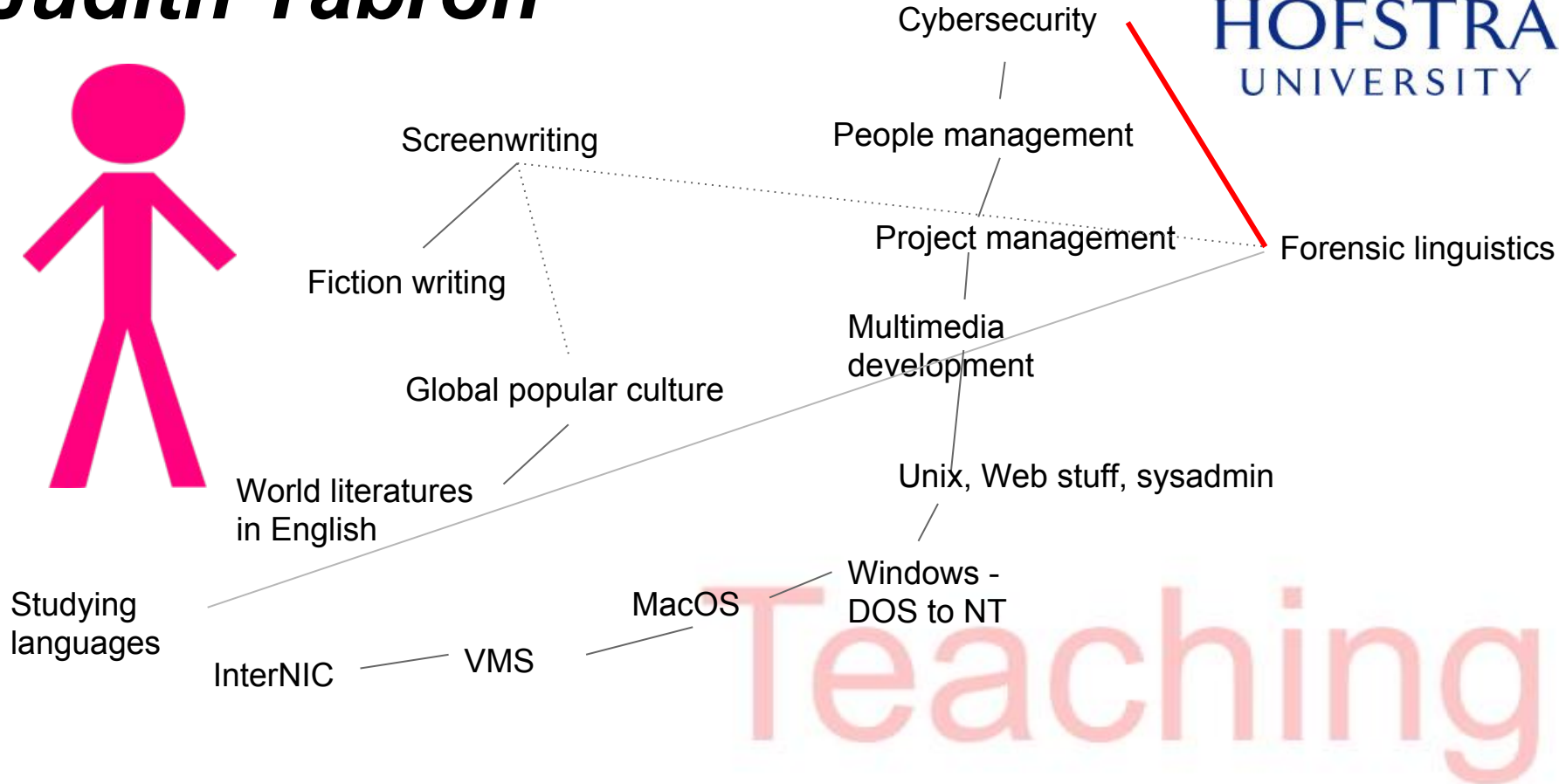
Forensic linguistics is the study of language as evidence for the law. It is a relatively new field and has not previously been applied to cybersecurity. Linguistic analysis uncovers several features of language interaction in a limited data set (recorded IRS phone scammers) that begin to answer how forensic linguistics could assist in cybersecurity defense.

Current phone security technology records all calls and in some cases analyzes them to ascertain the identity of the caller. Forensic linguistics could be used to **identify the crime, not the person**, and therefore has a broader - and cheaper - applicability.

Judith Tabron



HOFSTRA
UNIVERSITY



Teaching

Judith Tabron



tl;dr:



My mother was a sysadmin for DARPA.



HOFSTRA
UNIVERSITY

Cybersecurity and Forensic Linguistics: What's out there for securing the human?



HOFSTRA
UNIVERSITY

Phoneprinting: Pindrop

- **Won a patent and \$35million in investor funding last year**
- **At that time already used at 2 of the 4 biggest banks**
- **Builds a profile not just of your voice but of your voice on your phone line**

Phoneprinting: Pindrop



Phoneprinting: Pindrop



HOFSTRA
UNIVERSITY



Phoneprinting: Pindrop



HOFSTRA
UNIVERSITY



Identifies the
Caller
And the
Caller's
location



HOFSTRA
UNIVERSITY

Uptivity: “Stress detection”

Formerly CallCopy

- **Now available to companies “of all sizes”**
- **Sound based**
- **Keyword spotting**
- **Accuracy “greater than 80 to 90 percent”**
- **Real time**



HOFSTRA
UNIVERSITY

Uptivity: “Stress detection”

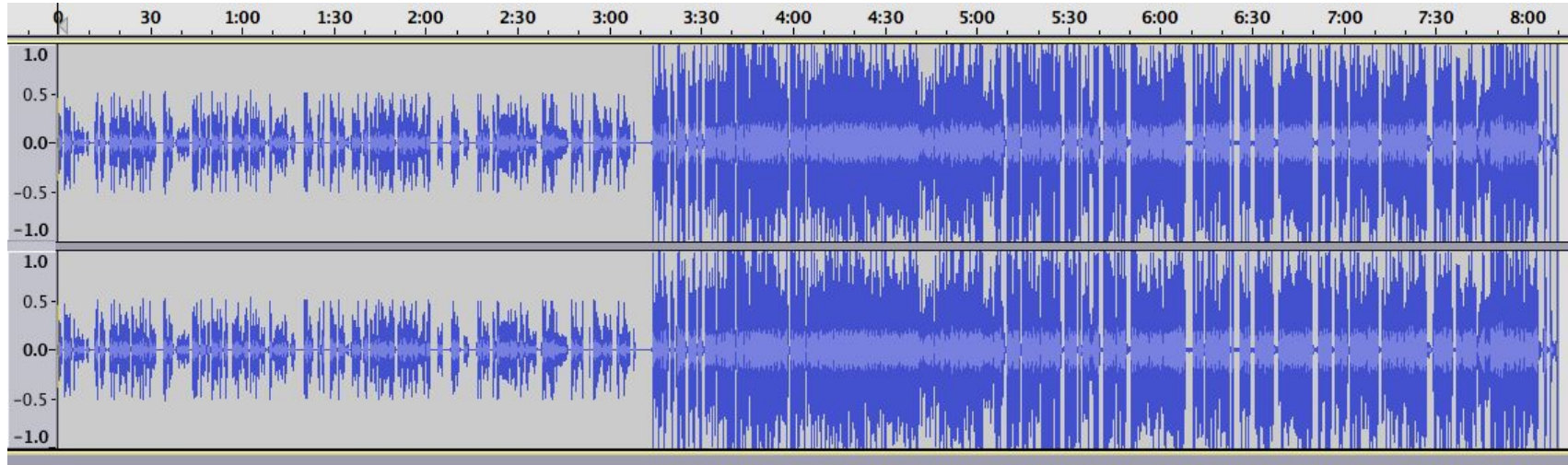
“Talk analysis can identify patterns within calls, such as long hold times or periods of silence, as well as the frequency of an agent cutting off a caller.” (*underlining mine*)

Identifies
keywords and
“emotion”
phonologically



HOFSTRA
UNIVERSITY

IRS scam phone call features both gaps and interruptions



Gaps

Interruptions

Competitors in this space



- **Verint** (acquired Victrio)
Creates a voiceprint for on-the-fly user identification
 - Identifies known criminals by voiceprint ¹
- **TrustId**
Verifies ANI before the phone call even starts - “pre-call authentication”
- **HP Autonomy Inc.**
Collects unstructured data for adding to CRMs, etc.

Wanted to improve phone security:

- Cheaper
- Fewer privacy issues
- Focus on a different goal
- Fail closed



Forensic Linguistics

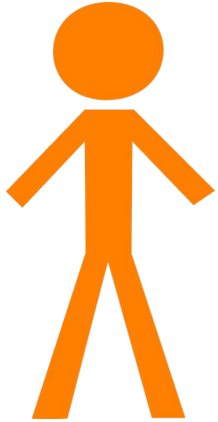
Applied linguistics to understand language
evidence (according to legal standards)

- Threat analysis
- Author attribution
- Legal language disputes
- Conviction exoneration (Innocence Project)
 ...not deception detection (for me)

Forensic *Linguistics*



HOFSTRA
UNIVERSITY



Sociolinguistics

Discourse analysis

“I’m sorry.” “Hey, it’s okay.”

Semantics

Pragmatics

“The dog ate my dessert.”



Morphology

eat, ate

Orthography

Phonology

/eit/

Forensic Linguistics



Roger Shuy and the “devil strip” case

-creating a profile to help locate the author of a ransom note

Forensic Linguistics



Do you ever want to see your precious little girl again? Put \$10,000 cash in a diaper bag. Put it in the green trash kan on the devil strip at corner 18th and Carlson. Don't bring anybody along. No kops!! Come alone! I'll be watching you all the time. Anyone with you, deal is off and dautter is dead!!!

Forensic Linguistics



*Do you ever want to see your precious little girl again? Put \$10,000 cash in a diaper bag. Put it in the green trash kan on the devil strip at corner 18th and Carlson. Don't bring anybody along. No kops!! Come alone! I'll be watching you all the time. Anyone with you, [the] deal is off and [your] **dautter** is dead!!!*

Forensic Linguistics



*Do you ever want to see your **precious** little girl again? Put **\$10,000** cash in a diaper bag. Put it in the green trash **kan** on the **devil strip** at corner 18th and **Carlson**.*

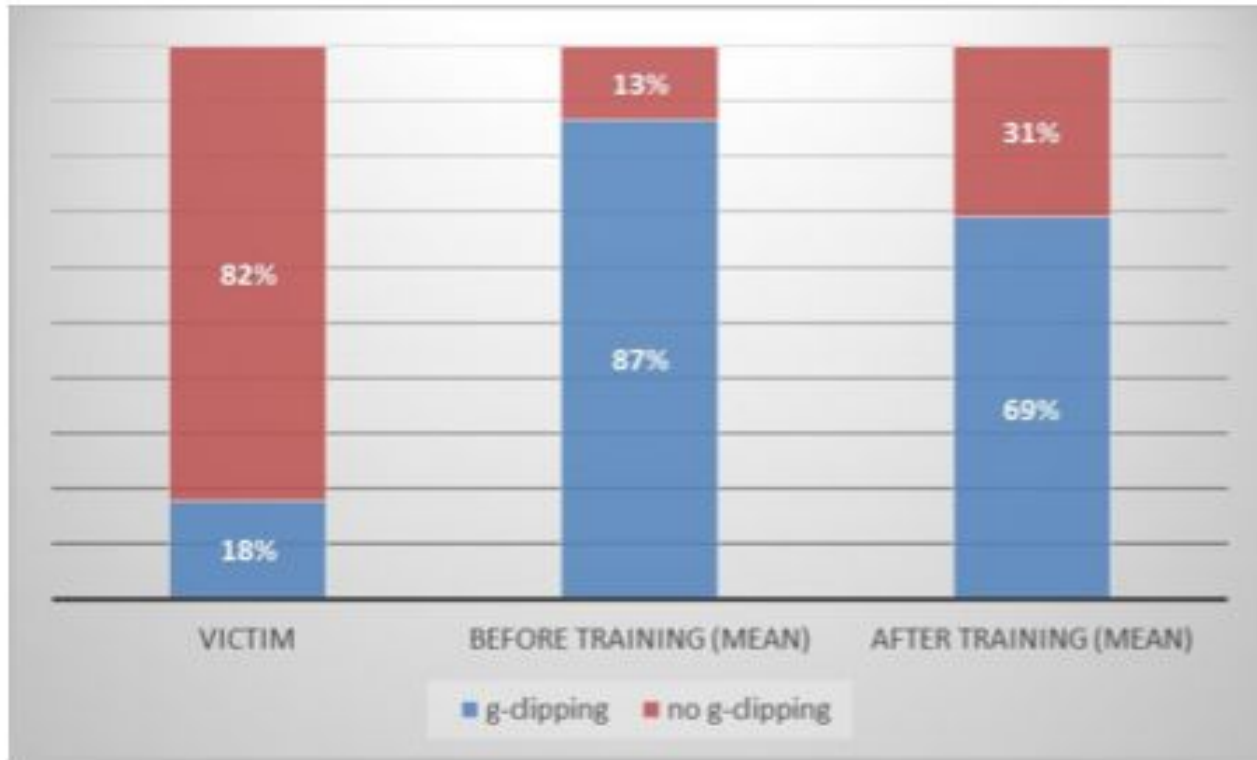
***Don't** bring anybody along.*

*No **kops**!! Come alone! I'll be watching you all the time. Anyone with you, deal is off and **dautter** is dead!!!*



HOFSTRA
UNIVERSITY

Training
police to
emulate the
language of
a potential
victim



G-clipping in Pilgrim data

Dr. Nicci MacLeod, Aston University, UK

Forensic Linguistics



Robert Leonard's work identifying the author of a threat letter that preceded a murder

Forensic Files episode 30 (season 11)

[“A Tight Leash”](#)

The author did not use negative contractions

Forensic Linguistics



Jim Fitzgerald of the FBI did similar profiling work --
the Unabomber case and with threat letters in
Australia

Forensic Linguistics



- Tammy Gales'** work on FBI's threat database -
What are the features of threats that are carried out versus threats that are not
- Corpus analysis

Forensic Linguistics



Tim Grant in the UK - work identifying the author of SMS messages in a murder case

- “TXT 4N6” article in various venues

Forensic Linguistics



My internship with Robert Leonard and Eric Freedman of the Hofstra Law School,
Distinguished Professor of Constitutional Rights

- Improving instructions for juries considering the death penalty

Forensic Linguistics



.... Not deception detection.

.... And not a “unique fingerprint”.

Forensic Linguistics

Can we identify features of social engineering phone interactions such that we could teach our community to recognize when they are occurring?



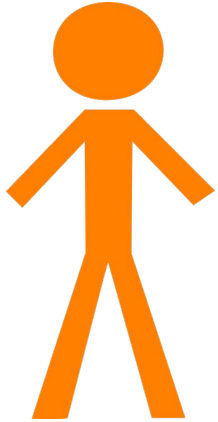
“Can you reset my password?”



Forensic *Linguistics*



HOFSTRA
UNIVERSITY



Sociolinguistics

“I’m sorry.” “Hey, it’s okay.”

Discourse analysis



Semantics

“The dog ate my dessert.”

Pragmatics

Morphology

eat, ate

Orthography

Phonology

/eit/

Forensic Linguistics

- Not automated
- Strengthens the human link, the weakest link in cybersecurity defense



“Can you reset my password?”



Forensic Linguistics

- Not automated
- Strengthens the human link, the weakest link in cybersecurity defense

Identifies the
Crime



“Can you reset my password?”



Forensic Linguistics



IRS Scam phone calls

“This January, the Treasury Inspector General for Tax Administration (TIGTA) announced they have received reports of roughly 896,000 contacts since October 2013 and have become aware of over 5,000 victims who have collectively paid over \$26.5 million as a result of the scam.”



“Police are coming to arrest you”



Forensic Linguistics

- Americans are uncertain about tax penalties
- And we tend to blame victims of fraud
- Fraud is underreported and underprosecuted





“Police are coming to arrest you”



Forensic Linguistics

Qualitative study of IRS scam phone calls
posted on YouTube

Transcribed and analyzed for *discourse*
features

 “I’m sorry.” “Hey, it’s okay.” 
~~“I’m sorry.” “Pass the bread.”~~

Forensic Linguistics



Polar tag questions, okay?

Forensic Linguistics

3:25	50		All right ma'am you would be getting this tax pay order (.) there, OK?
3:30	51	OK.	
3:31	52		So, ma'am but first we have to go to that store before going to that store ma'am, because the store people they won't be, they don-, they just only accept the cash, so you have to go to your bank first, need to withdraw the money, and then you have to go to the store, to purchase this tax pay vouchers. All right?

Forensic Linguistics



Polar tag questions:

7 to 33 occurrences
on the part of the
scammers in my
samples

Prefer a response

Prefer a positive response

Prefer a response that accepts the terms of the
previous statement

Forensic Linguistics



Topic control & Question deferral

“Can I just --”

“No.”

Forensic Linguistics



HOFSTRA
UNIVERSITY

Speaker B informational questions and Speaker C responses

Question deferral

who do I send them to	deflected to future
what's your address	deflected to future
why don't you just give me, um, your address	deflected to future
so can you	interrupted
can you let me know what I do	deflected with "anger"
should I just come home and call you	deflected with "anger"
who to send this to would you	deflected to future
where do I mail this	deflected to future
do I mail it to the IRS	deflected to future
Are you ready?	responds with a question
are you ready? [2 nd occurrence]	responds with a question
do you want me to put you on hold...	deflected to future
which Winn Dixie do you want me to go to?	deflected to future
What's your name?	ignored
Are you gonna hold?	Answered ("Yes")

Forensic Linguistics



Violations of narrative structure

Forensic Linguistics

Violations of narrative structure

William Labov, 1960s, ff

1. Abstract: What is the story about?
2. Orientation: Who, when where, how?
3. Complicating action: Delays resolution, adds suspense...
4. Evaluation: Who's the hero? Who's the bad guy?
5. Result / Resolution: What happened in the end?
6. Coda "And that's why I can never go back there again."

Forensic Linguistics



Violations of narrative structure

“You owe the IRS.

There’s nothing you can do about it.

Police are coming to arrest you.

”

.....

Forensic Linguistics

Violations of narrative structure



May be the toughest to realize in the moment

Tends to require after-the-fact analysis
... but I'm giving you the information

Forensic Linguistics

Polar tag questions & question deferral:

Can you strengthen your organizations
by
educating them about these linguistic
peculiarities?



Forensic Linguistics



**The Institute for Forensic Linguistics, Threat
Assessment, and Strategic Analysis at Hofstra
University**

Robert Leonard
Tammy Gales



Forensic Linguistics

Judith Tabron

judith.tabron@gmail.com

@judithtabron



HOFSTRA
UNIVERSITY

