ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 13, No. 8, pp. 2116-2122, August 2023
DOI: https://doi.org/10.17507/tpls.1308.28

The Speech of Social Media Influencers in Najd: Introducing a New Source of Sociolinguistic Data

Nasser M. Alajmi
Department of English, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia

***Abstract*—This study explores using the speech of Najdi social media influencers in Snapchat and TikTok as a source of sociolinguistic production data. Saudis in general, Najdis in particular, post their daily life vlogs on these apps building a huge volume of recorded speech in the local dialects. This type of content has been examined to introduce the possibility of using it as a source of production data alongside/instead of the sociolinguistic interview. The speech of 18 social media influencers, representing three sub-dialects of Najd, has been transcribed and analyzed to test whether it is natural and not mostly formal. The variables have been ranked according to Labov's (1972) classification of speakers' awareness: stereotype, marker and indicator. It was found in this study that, similar to data elicited from the interview, social media influencers tend to use their local variants which are classified as markers and/or indicators, and they avoid using stereotypes. In some social settings, the social media influencers were even more spontaneous than they would be in an interview. The study suggests that the speech of social media influencers in Najd is an easily accessible, larger, and better-quality source of sociolinguistic production data.**

***Index Terms*—indicator, marker, production data, stereotype**

I. Introduction

The bulk of production data (recorded speech) in sociolinguistics has been collected via the sociolinguistic interview. This method has evolved and been adjusted to accommodate different sociolinguistic settings. For instance, the Labovian framework of the interview which elicits casual/careful styles (Labov, 1972) does not work for other languages, e.g., Arabic because it will trigger the use of Standard Arabic rather than different styles (Al-Wer, 2013). As technology advances, it can be observed that sociolinguistic studies relying on social media data are gaining traction (Sun et al., 2021). In this study, the speech of social media influencers in Najd will be analyzed sociolinguistically to further explore the possibility of using it as a reliable data source for Najdi dialects. In the remainder of this section, we shall shed some light on social media evolution in Saudi Arabia, the dialects under study, and the linguistic variables.

A. Social Media Evolution in Saudi Arabia

According to Helmond and Van der Vlist (2019), in the early 2010s, social media platforms boomed across the world, and the number of their users is rapidly increasing every year. In the past 7 years, the competition between tech giants on who will throne the world was fierce. Several platforms were launched and only a few have succeeded/survived. It is no surprise that social media platforms have attracted users across the world at a considerably different rate. A social media platform could be popular in one country but not necessarily in others. Facebook for example is not as popular in Saudi Arabia as it is in Egypt. On the other hand, Twitter is more popular in Saudi Arabia than it is in Egypt.

Saudi Arabia is one of the countries which adopted social media massively. In January 2021, the number of active social media users in Saudi Arabia was 79.3% of the total population, compared to Egypt which was 49% (Kemp, 2022). As in any part of the world, social media platforms in Saudi Arabia become popular, and then they fall out of trend as new appealing ones emerge. The top five social media platforms in Saudi Arabia are listed in Table 1. In this study, we shall focus on the last two platforms in the table, Snapchat and TikTok, due to the type of content they provide.

Table 1
Top Five Social Media Platforms in Saudi Arabia (the Global Statistics, 2022)

Social media platform	Percentage	Number of Users
WhatsApp	87.40%	30.45 M
Instagram	78.10%	27.21 M
Twitter	71.90%	25.05 M
Snapchat	68.80%	23.97 M
TikTok	63.60%	22.16 M

Snapchat started as a messenger app that allows users to send pictures and videos to be viewed once. It later evolved to allow individuals to post stories, which can be viewed within 24 hours. TikTok then followed the steps of Snapchat by allowing users to post videos, but it added the feature of live streaming. Saudis have begun to post their daily routines in the form of successive short videos (stories or vlogs). The content of these stories includes cooking, gardening, traveling, hunting, fashion, life coaching, etc., but the main genre is *lifestyle*. It was not long before

advertisers realized that these platforms are not only cheaper than mainstream channels but even better yielding. This was the catalyst that caused this type of content to boom. Therefore, there are countless hours of recorded speech in spoken Najdi dialects available online. From an academic perspective, this content may well be regarded as a gold mine for sociolinguistics.

B. Najdi Dialects

It is proposed that the most conservative contemporary spoken Arabic variety is Najdi, and this is attributed to the fact that Najd is relatively geographically isolated (Ingham, 1994, p. 5; Versteegh, 2001, p. 193). There are many conservative features that Najdi Arabic, along with other Arabic varieties, has preserved, e.g., the retention of interdentals and dual noun suffixes. However, Najdi Arabic, unlike other dialects, has retained archaic features from Classical Arabic. These Archaic features are the internal passive as in for example *yaktib* 'writing' > *yiktab* 'being written' and indefinite suffix *tanwīn -in* as in *kitab-in* 'a book' (Ingham, 1994; Palva, 2006).

Ingham (1994, p. 5) classifies the dialects of Najd as the following regional subgrouping:

1. Central Najdi: The dialects of bedouin tribes and the sedentary population in the central region.
2. Northern Najdi: The dialect of Shammar tribe and the northern parts of the peninsula.
3. Mixed Northern-Central: The dialect of Qassim region.

Although these dialects share Najdi features with one another as opposed to Western (Hijaz) or Eastern dialects in the Arabian Peninsula, there are some minor differences, mainly in the object suffixed pronouns (henceforth, SPs). More on these will be highlighted in the next section. Administratively, the dialects of Najd are spoken in the provinces of Riyadh, Qassim, and Hayil (see Map 1).

Figure 1. Map of the Administrative Provinces of Saudi Arabia

C. Linguistic Variables

The linguistic variables in this study will be presented by dialect group (Central, Northern, Mixed Northern-Central). The variables will also be classified according to Labov's (1972) taxonomy of speakers' awareness of a linguistic feature: stereotype, marker, and indicator. According to Labov (1972, p. 314), an indicator is a linguistic feature that shows no style shifting and tends to have little evaluation force. Markers, on the other hand, show both stylistic and social stratification. The highest in term of speakers' awareness is a stereotype, which is socially marked by society. The collection of variables is mainly obtained from previous studies. However, some linguistic features along with their classification according to speakers' awareness are derived from the researcher's knowledge of Najdi dialects (being a native speaker of Najdi Arabic) and the help of 12 informants (4 from each dialect group). The sociolinguistic salience of a variable could be measured in several ways (see Alajmi & Alghannam, 2022). In this study, however, we will rely on informants from each dialect, as it is one of the most reliable methods. Their classification will be treated as provisional until it agrees with the sociolinguistic patterns in the data, as we shall see below. The linguistic variables in this study and their classification into stereotype (S), marker (M), or indicator (I) are provided in Table 2 below.

Northern Najdi Dialect: this dialect, as stated above is spoken in the province of Hayil. According to Ingham (1994, p. 193), the realization of some of the object suffixed pronouns (henceforth, SPs) are shared with Mixed Northern-Central (henceforth, Qassimi dialect), but not with Central Najdi dialect, see Table 2. An additional variable that Northern Dialect shares with Qassimi dialect is changing *-a:y-* to *-e:-* as in *ne:m* < *na:yim* 'asleep'. A linguistic feature, specific to Northern Dialect, is the realization of the feminine marker *-ah/-at* as *-ay*, e.g., *fazʕ-ay* < *fazʕ-ah* 'help-FEM.'. Another interesting feature of the Northern Dialect is the realization of one of the most frequent words *ʃay* 'thing' as *ʃin*.

Qassimi dialect: besides the SPs that Qassimi dialect shares with Northern Dialect only, there are a few other variables that are specific to this dialect. One feature is the realization of /a/ in the feminine suffix *-ah* (not *-at*) as *-ʉh*, e.g., *fazʕ-ʉh* < *fazʕ-ah* 'help-FEM.' and *kis-ʉh* < *kis-ah* 'bag-FEM.'.

Central Bedouin/Sedentary: The Central Dialect is spoken by Bedouin tribes and sedentary groups in what is known today as Riyadh province. The Bedouin tribes speak a shared dialect, but it can be further subdivided into South-Western Bedouin and Western Bedouin, each of which descends from a different lineage and was known to inhabit different areas of the Arabian Peninsula (Alajmi, 2019). In this study, however, the distinctive features of each bedouin tribe will be highlighted rather than treating the Bedouin dialect as a single entity. This is because some of the marked features in the Central Dialect are restricted to a particular tribe. As for the sedentary groups, the dialect will be based on the cities within Riyadh province in which the sedentary groups are known to live. This is because the history and social structure of the sedentary population are different than those of the bedouins (for further details see Alajmi, 2019, p. 86). The first linguistic variable is evident in the speech of the Ajmi tribe. The realization of the 2.S.F SP is *-ʃ* < *-ik*, as in *kita:b-ʃ* 'your book-F'. The second distinctive feature is the insertion of *ʔana* '1.S.C' in perfective verbs by Dosari tribe (Ad-Darsoni, 2013). The perfective verb *xaðe:tha* 'I took it.F' is realized in Dosari dialect as *xaðe:t-ana-ha*. The third variable is the lowering of /i/ and /e:/ to /a:/ by Otaibi tribe, as in *ba:ðˁ* < *be:ðˁ* 'eggs' and *ya-ba:* < *ya-bi* 'he wants'. Another feature by Otaibi tribe is the realization of the 3.S.M demonstrative pronoun *haðəh* < *ha:ða* 'this'. The fifth feature is the addition of the case marking *fathah* in 2.S.M perfective verb conjugation *-t* as *-ta* in the sedentary dialect of Sudair and Al-Washim (Ingham, 1994, p. 194), e.g., *xaðe:-ta* < *xaðe:-t*. This feature, however, is attested in the speech of sedentary groups in other cities, as observed by the researcher. The last sedentary feature is the drop of the epenthetic vowel in the 2.S.M suffixed possessive pronoun *-ik* as in *siyya:rat-k* < *sayya:rat-ik* 'your car'.

TABLE 2
LIST OF THE LINGUISTIC VARIABLES

Linguistic Variables		Najdi Dialect						Classification Based on Speakers' awareness
		Northern Dialect	Qassim Dialect	Central Dialect				
				Ajmi	Otaibi	Dosari	Sedentary	
SPs	'me'	*-an*	*-an*	*-ni*	*-ni*	*-ni*	*-ni*	M / I
	'your.s.m'	*-ak*	*-ik*	*-ik*	*-ik*	*-ik*	*-ik*	I
	'your.s.f'	*-k / -ts*	*-k / -ts*	*-ʃ*	*-k*	*-k / -ts*	*-k / -s*	S
	'you.pl'	*-kam*	*-kum*	*-kum*	*-kum*	*-kum*	*-kum*	I
	'them'	*-ham*	*-hum*	*-hum*	*-hum*	*-hum*	*-hum*	I
	'him'	*-uh/w*	*-uh/w*	*-ih/h*	*-ih/h*	*-ih/h*	*-ah/h*	S / M
	'her'	*-ah*	*-ah*	*-ha*	*-ha*	*-ha*	*-ha*	S / M
-a:y- > *-e:-*		✓	✓	×	×	×	×	S
F marker *-ah/-at* > *-ay*		✓	×	×	×	×	×	S
ʃay 'thing' > *ʃin*		✓	×	×	×	×	×	S / M
F marker *-ah* > -uh		×	✓	×	×	×	×	M / I
insertion of *ʔana*		×	×	×	×	✓	×	S
/e:/ > /a:/		×	×	×	✓	×	×	S
3.s.m demonstrative pronoun		*ha:ða*	*ha:ða*	*ha:ða*	*ha:ðah*	*ha:ða*	*ha:ða*	S / M
Verb conjugation 'you.s.m'		*-t*	*-t*	*-t*	*-t*	*-t*	*-ta*	S
vowel drop (*-ik* > *-k*)		×	×	×	×	×	✓	I

The linguistic variables in Table 2 are listed according to the chronological order in which they were discussed above, except for two SPs of Ajmi and sedentary dialect which were listed with the relevant group (SPs). Both marked and standard variants were provided for the sake of comparison. Each marked variant is classified in terms of speakers' awareness of said variant. The classifications, as stated above, are based on the researcher's knowledge of the dialects in addition to consultation with 12 informants. For some variants, there was disagreement on whether a variant should be classified as, for example, a stereotype or a marker. Hence, all classifications were provided in the table.

D. The Study

This study aims to further explore social media as a viable source of sociolinguistic data. The circumstances in Saudi Arabia have led social media platforms to thrive. Thousands of hours of recorded speech in colloquial dialects can be observed and investigated. Labov's (1984) argument of the observer's paradox and how we can never be able to obtain authentic data on a vernacular has already been proven true. In a sociolinguistic interview, no matter how an interviewee is relaxed, he/she may never produce particular linguistic features except with family and close friends, i.e., stereotypes. Since speakers are expected to avoid stereotypes and will probably use markers and indicators in a traditional interview, it is hypothesized after initial observations of social media influencers' speech that this is also the case. Therefore, the setting of the sociolinguistic interview is postulated to yield the same type of data as vlogs by influencers. It is even argued that in some settings in such vlogs, e.g., group chat with close friends, speakers tend to be more relaxed than in an interview. If the hypothesis that states that social media data is similar to or better than data elicited from the sociolinguistic interview is true, it can be stated that a new era of sociolinguistic data collection in Najd has begun. The data in social media are endless and could be collected with less time, effort, and expenses. We aim to answer the following research questions:

- What are the sociolinguistic patterns in the speech of social media influencers?
- To what extent could the speech of social media influencers be used in future sociolinguistic studies?

II. Literature Review

As stated earlier, linguistic studies on social media are attracting the attention of researchers in the field. According to Sun et al. (2021), the majority of linguistic studies on social media are in the sub-fields of sociolinguistics, discourse analysis, pragmatics, and applied linguistics. It was also found that most studies are conducted on the social media platforms of Twitter and Facebook.

Reviewing studies on social media in other strands of linguistics is beyond the scope of this study. We shall focus on only sociolinguistic studies on social media. Most studies on language variation in social media rely on data retrieved from written content on Twitter and Facebook. Examples of these studies are: Aboh and Ezeudo (2020); Almeman and Lee (2013); Bazarova et al. (2013); Carr et al. (2012); Ceron and D'Adda (2016); Christiansen (2019); McDonnell (2020); Mubarak and Darwish (2014); Tankosić and Dovchin (2021). In these studies, the analysis and mapping of features are based on both linguistic and non-linguistic data. An example of non-linguistic data is the geo-location from which a tweet or Facebook comment has originated. In such studies, lexical, morphological, and grammatical variables can be explored, but not commonly phonological. However, in some studies, e.g., Eisenstein (2015) and Doyle (2014), it is assumed that the phonological variation will be represented in the orthography (i.e., tweets or comments). In the case of the Arabic language, this assumption is based on the fact that most social media users will intentionally use unconventional orthographical norms (i.e., deviate from Standard Arabic) to exhibit phonological dialectal differences, as in Alshutayri and Atwell (2021). For example, it is presumed that a Kuwaiti speaker will write *dija:jah* (دياية) 'a chicken' instead of *didʒa:dʒah* (دجاجة). Nevertheless, such studies can be valuable for regional variations that are highly contrastive and can possibly appear in the orthography. In other words, a regional linguistic variable featuring two consonantal variants, e.g., Baghdadi Arabic /q/ > /k/ as in *kalbi* 'my heart' (Abu-Haidar, 2006) can be represented in the orthography as opposed to most variables in this study, which are short vowels. This is mainly because short vowels in Arabic are represented by diacritics (Watson, 2002) which are absent in almost all written content on social media platforms. The type of social media data in which such variables can be examined is recorded speech in the form of either audio or video.

There are few studies that examined recorded speech in social media, e.g., Androutsopoulos and Ziegler (2004); Biel and Gatica-Perez (2010); Concha (2019); Sutrisno and Ariesta (2019). However, to my knowledge, no study has employed social media spoken content as a substitute for the sociolinguistic interview to elicit production data, at least for Najdi dialects.

III. Methodology and Data Collection

The data in this study was collected from social media influencers' stories (short videos that are posted and can be viewed within 24 hours) on Snapchat and TikTok. The sample is comprised of 18 social media influencers: 5 representing Northern Dialect, 5 representing Qassimi Dialect, and 8 representing Central Dialect (2 for each sub-group). The sample is selected carefully, taking into consideration the following factors: (a) speakers are roughly the same age, (b) speakers have roughly the same level of education (college degree), (c) speakers who do not live in the same area they represent or provide specialized content (e.g., tech advice) are excluded but observed. The researcher listened to and transcribed at least four hours of speech for each speaker, monitoring every social setting. The data was analyzed quantitatively using the *principle of accountability*, quantifying all occurrences of a linguistic variable including zero variant (Labov, 1972).

IV. Results and Discussion

The results for each dialect group (Northern Dialect, Qassim Dialect, and Central Dialect) are provided in Tables 3, 4, and 5, respectively.

Table 3
Frequency of Variants for Northern Dialect

Variable	Target variant	Other variant	Frequency	%	Classification
'me' SP	*-an*	*-ni*	88/95	92.63	M / I
'your.s.m' SP	*-ak*	*-ik*	73/75	97.33	I
'him' SP	*-uh/w*	*-ih/h /-ah/h*	122/131	93.12	S / M
'her' SP	*-ah*	*-ha*	81/86	94.18	S / M
-a:y- > -e:-	✓	×	5/16	31.25	S
'you.pl' SP	*-kam*	*-kum*	76/77	98.70	I
'them' SP	*-ham*	*-hum*	63/67	94.02	I
F marker *-ah/-at > -ay*	✓	×	4/23	17.39	S
ʃay 'thing' > *ʃin*	✓	×	2/15	13.33	S / M

TABLE 4
FREQUENCY OF VARIANTS FOR QASSIM DIALECT

Variable	Target variant	Other variant	Frequency	%	Classification
'me' SP	*-an*	*-ni*	164/181	90.6	M / I
'him' SP	*-uh/w*	*-ih/h /-ah/h*	179/190	94.21	S / M
'her' SP	*-ah*	*-ha*	124/139	89.20	S / M
-a:y- > *-e:-*	✓	×	8/31	25.80	S
F marker *-ah* > -uh	✓	×	145/152	95.39	M / I

TABLE 5
FREQUENCY OF VARIANTS FOR CENTRAL DIALECT

Sub-group	Variable	Target variant	Other variant	Frequency	%	Classification
Sedentary	Verb conjugation 'you.S.M'	*-ta*	*-t*	4/90	4.4	S
	vowel drop (*-ik* > *-k*)	✓	×	98/103	95.14	I
Ajmi	'your.s.f' SP	*-ʃ*	*-ik*	1/57	1.75	S
Otaibi	/e:/ > /a:/	✓	×	2/35	5.71	S
	3.S.M demonstrative	*ha:ðah*	*ha:ða*	26/44	59.09	S / M
Dosari	insertion of *ʔana*	✓	×	2/79	2.53	S

The results show that SPs, most of which are ranked as M/I, show high usage of the variant of each dialect group. As pointed out in Alajmi (2019, p. 186), the morphophonological variables in Najd in general, SPs in particular, are not as salient as phonological variables. Najdi speakers tend to use the variant they grow up hearing and they don't tend to avoid it as long as it does not attract negative feedback. Although some SPs are ranked as S/M by informants, the data show that although speakers are aware of it on a high level, it is still not a "stereotype". For example, the Qassimi Dialect *-uh* 'him' and *-ah* 'her' are some of the most, if not the most, salient variants of Qassimi Dialect; yet they are used frequently in almost all social settings. This, however, can be attributed to the high level of solidarity and pride in their identity. On the other hand, the Ajmi SP *-ʃ* 'your.s.f' has almost disappeared from Ajmi Dialect. This is not attributed to weak identity, but rather to the number of users of the other variant *-ik* (being the majority) and probably due to the high level of phonological contrast between the variants. As for the rest of the variables, the rate of usage of the target variant agrees with the level of speakers' awareness: stereotype variants are avoided while indicators are used frequently by the relevant group.

It should be noted that the social setting in which the videos are recorded has played a major role in the style of influencers. If they speak about a donation campaign (hence religious context), most of them tend to switch to Standard Arabic. On the other hand, the target variants are most reported when they are speaking casually with friends who speak the same dialect. Lastly, the type of content the influencer is producing also plays a role in determining the frequency of using either the target or the standard variant. Social media influencers who categorize their content as 'lifestyle influencers' – which includes cooking, travelling, farming, and comic sketches – tend to be the most natural, i.e., use their target variants. on the other hand, those who provide specialized content, e.g., tech advice and life coaching, tend to use either Standard Arabic or the standard variants of all dialect groups.

It can be stated that the speech of Najdi social media influencers is a viable source of sociolinguistic production data. This is because the sociolinguistic behavior of speakers is not different from that of the interview. In fact, the researcher has noted that when interviewing speakers who speak a different sub-dialect of Najd, they tend to use formal language throughout most parts of the interview. On the contrary, social media influencers who were acquainted with the platform and their audience are likely to produce more casual forms.

V. CONCLUSION

The collection of data in sociolinguistics has been problematic as it is difficult to elicit the most casual form of language 'the vernacular'. The dilemma is that "any systematic observation of a speaker defines a formal context where more than minimum attention is paid to speech …[however] face-to-face interviews are the only means of obtaining the volume and quality of recorded speech that is needed for quantitative analysis" (Labov, 1984, p. 28). However, the data in this study which was collected from the speech of social media influencers in Najd suggests that in this new type of data speakers tend to be more relaxed, hence use the local features. This indicates that at least in the case of Najdi Arabic, sociolinguistic production data can be collected from social media platforms that are used to post daily vlogs. Not only is this type of data abundant and easily accessible, but the authenticity of speech is arguably better than what is achieved by the traditional method.

ACKNOWLEDGMENTS

The author extends his appreciation to Prince Sattam Bin Abdulaziz University for funding this research work through the project number (PSAU/2022/02/21751)

REFERENCES

[1] Aboh, S. C., & Ezeudo, O. C. (2020). Interactions on Facebook and Twitter: A communicative action perspective. *Theory and Practice in Language Studies*, *10*(11), 1351-1358. https://doi.org/10.17507/tpls.1011.02

[2] Abu-Haidar, F. (2006). Baghdad Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 222- 231). Leiden: Brill.

[3] Ad-Darsoni, S. (2013). *Mu'jam allahajaat almahkiyyah fi almamlakah al'arabiyah assi'udiyah*. Jazan University. Jazan, Saudi Arabia.

[4] Alajmi, N. (2019). *The Bedouin-Sedentary Dichotomy in Najd: A Sociolinguistic Study* [Unpublished doctoral dissertation]. University of York, UK.

[5] Alajmi, N. M., & Alghannam, A. G. (2022). The sociolinguistic salience of linguistic variables in Najdi Arabic. *World Journal of English Language*, *12*(6), 114-128. https://doi.org/10.5430/wjel.v12n6p114

[6] Almeman, K., & Lee, M. (2013), Automatic building of Arabic multi dialect text corpora by bootstrapping dialect words, In *1st International Conference Communications, Signal Processing, and their Applications* (ICCSPA), Sharjah, UAE, IEEE.

[7] Alshutayri, A., & Atwell, Eric. (2021). Classifying Arabic dialect text in the Social Media Arabic Dialect Corpus (SMADC). *Proc. of the 3rd Workshop on Arabic Corpus Linguistics*, pp. 51-59.

[8] Al-Wer, E. (2013). Sociolinguistics. In J. Owens (ed.), *The Oxford Handbook of Arabic Linguistics*. Oxford: Oxford University Press, pp. 241-263

[9] Androutsopoulous, J., & Ziegler, E. (2004). Exploring language variation on the Internet: Regional speech in a chat community. In Gunnarsson, B.-L. et al. (eds.), *Papers from ICLaVE 2 Uppsala Papers from the Second International Conference on Language Variation in Europe*, ICLaVE 2. Uppsala, Sweden: Department of Scandinavian Languages, Uppsala University

[10] Bazarova, N. N., Taft, J. G., Choi, Y. H., & Cosley, D. (2012). Managing impressions and relationships on Facebook. *Journal of Language and Social Psychology*, *32*(2), 121-141. https://doi.org/10.1177/0261927x12456384

[11] Biel, J., & Gatica-Perez, D. (2010). Voices of vlogging, in *ICWSM 2010 - Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, pp. 211-214.

[12] Carr, C. T., Schrock, D. B., & Dauterman, P. (2012). Speech acts within Facebook status messages. *Journal of Language and Social Psychology*, *31*(2), 176-196. https://doi.org/10.1177/0261927x12438535

[13] Ceron, A., & D'Adda, G. (2016). E-campaigning on Twitter: The effectiveness of distributive promises and negative campaign in the 2013 Italian election. *New Media & Society*, *18*(9), 1935-1955. https://doi.org/10.1177/1461444815571915

[14] Christiansen, M. S. (2019). "Listisimo para los #XVdeRubi:" constructing a chronotope as a shared imagined experience in Twitter to enact Mexicanness outside of Mexico. *Lingua*, *225*, 1-15. https://doi.org/10.1016/j.lingua.2019.05.002

[15] Concha, S. (2019). Analysis of Strategic and Sociolinguistic Competences Elicited Through Online Chat Activities. *Praxis Pedagógica*, *19*(24), 41-59. http://dx.doi.org/10.26620/uniminuto. praxis.19.24.2019.41-59

[16] Doyle, G. (2014). Mapping dialectal variation by querying social media. *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics.* https://doi.org/10.3115/v1/e14-1011

[17] Eisenstein, J. (2015). Systematic patterning in phonologically-motivated orthographic variation. *Journal of Sociolinguistics*, *19*(2), 161-188. https://doi.org/10.1111/josl.12119

[18] Helmond, A., & Van der Vlist, F. N. (2019). Social media and platform historiography: Challenges and opportunities. *TMG Journal for Media History*, *22*(1), 6-34. https://doi.org/10.18146/tmg.434

[19] Ingham, B. (1994). *Najdi Arabic: Central Arabian*. Amsterdam: John Benjamins Publishing Company.

[20] Kemp, S. (2022). Digital 2022: Egypt. *Datareportal*. Retrieved March 10, 2023, from https://datareportal.com/reports/digitalegypt.

[21] Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia, PA: University of Pennsylvania Press.

[22] Labov, W. (1984). Field methods of the Project on Linguistic Change and Variation. In J. Baugh, & J. Sherzer (Eds.), *Language in use: Readings in sociolinguistics* (pp. 28-53). Englewood Cliffs: Prentice Hall.

[23] McDonnell, A. (2020). Clinton stated, Trump exclaimed! Gendered language on Twitter during the 2016 presidential debates. *Journal of Language and Politics*, *19*(1), 71-88. https://doi.org/10.1075/jlp.19085.mcd

[24] Mubarak, H., & Darwish, K. (2014), Using Twitter to collect a multi-dialectal corpus of Arabic, In *Proceedings of the EMNLP, Workshop on Arabic Natural Language Processing* (ANLP), Doha, Qatar, Association for Computational Linguistics, pp.1-7.

[25] Palva, H. (2006). Dialects: Classification. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, & Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 604-613). Leiden: Brill.

[26] Saudi Arabia Social Media Statistics 2022: Most popular platforms. *The Global Statistics - The Data Experts | Statistical Data Reports.* (2022, August 30). Retrieved September 10, 2022, from https://www.theglobalstatistics.com/saudi-arabia-social-media-users/

[27] Sun, Y., Wang, G., & Feng, H. (2021). Linguistic studies on social media: A bibliometric analysis. *SAGE Open*, *11*(3), 1-12 https://doi.org/10.1177/21582440211047572.

[28] Sutrisno, B., & Ariesta, Y. (2019). Beyond the use of code mixing by social media influencers in Instagram. *Advances in Language and Literary Studies*, *10*(6), 143-151. https://doi.org/10.7575/aiac.alls.v.10n.6p.143

[29] Tankosić, A., & Dovchin, S. (2021). The impact of social media in the sociolinguistic practices of the peripheral post-socialist contexts. *International Journal of Multilingualism*, 1-22. https://doi.org/10.1080/14790718.2021.1917582

[30] Versteegh, K. (2001). *The Arabic language*. Edinburgh, UK: Edinburgh University Press.

[31] Watson, J. (2002). *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Nasser M. Alajmi did his BA in English Language from King Saud University and graduated magna cum laude in 2007. In 2014, he got his MA from California State University Fullerton in Linguistics. In 2019, he was awarded his Ph.D. in linguistics from the University of York, United Kingdom. His Ph.D. supervisor was the prominent figure of sociolinguistics (language variation and change) Paul Kerswill. The Ph.D. thesis investigated the sociolinguistics of Najdi Arabic.