

„De ha a gondolat megrontja a nyelvet,
akkor a nyelv is megronthatja a gondolatot.”
George Orwell: *A politika és az angol nyelv*

Adattengerbe merülve

Nyelvhasználatunk, ezernyi egyéb funkciója mellett, diagnosztikai információkat hordozhat mentális állapotunkról. E triviálisnak tűnő megállapítás mögött az összefüggések számos rétege tárható fel. Mert pszichopatológiai értelemben egyáltalán nem mindegy, milyen arányban használunk múlt idejű igéket, tagadó formulákat, általános és birtokos névmásokat vagy éppen általunk kreált "neologizmákat". Hogy mennyire koherens az általunk létrehozott szöveg, mennyire hatják át érzelmeiket kifejező szavak. És így tovább.

A modern pszicholingvisztika egyik leggyorsabban fejlődő területe éppen a mentális kórképek szövegek közötti azonosítása. Vagyis olyan nyelvi elemek meghatározása, amelyek más nyelvi elemekhez képest nagyobb megbízhatósággal jelzik előre egy-egy meghatározott pszichés zavar jelenlétét. Ahogyan ez a terület egyre inkább elfogadottá vált a kvantitatív módszerek által dominált tudományos világban, fontos szerepet játszott a *big data* korszak térnyerése, és ezzel együtt az online térben gyarapodó szöveganyag mérhetetlen gazdagodása. A világ lakosságának több mint fele (4,2 milliárd ember) aktívan posztol a közösségi média oldalakon, átlagosan napi két és fél órát töltve ezekben az online terekben (We Are Social és Hootsuite, 2021), nem túlzás tehát azt állítani, hogy az elmúlt években szinte felfoghatatlan méretűvé dagadt az elemezhető szövegtömeg (pl. közösségi média bejegyzések, posztok, tweetek stb. formájában). Nem csoda, hogy a statisztikában jártasabb kutatók (leginkább a *gépi tanulás* modelljeit alkalmazva, vagyis megadott szabályok alapján öntanulásra képes algoritmusokat tesztelve) hamar belevetették magukat a végelethetetlen adattengerbe. A kutatási terület/módszer gyűjtőneve lett a szövegbányászat (*text mining*). A manuálisan belátható időn belül nem elemezhető vagy kódolható adattömeg az automatizált modelleknek köszönhetően olyan szabályszerűségekre világított rá, amelyek a nyelvhasználaton alapuló pszichodiagnosztika új fejezetét nyitották meg.

Nézzünk egy szemléletes példát: a depresszió nyelvi indikátorainak szövegbányászati feltárását! Az elmúlt években publikált tanulmányok (pl. Eichstaedt és mtsai., 2018; CACHEDA és mtsai., 2019; Trifan és mtsai., 2020) első lépésként azonosították azokat a Facebook, Twitter vagy Reddit felhasználókat, akiket korábban depresszióval diagnosztizáltak (olyan posztok alapján, mint: „2020-ban kaptam depresszió diagnózist”, „depresszióval kezelnek”, „antidepresszánt szedek”). A közösségi média bejegyzések lingvisztikai elemzését követően különbséget találtak a depressziós felhasználók és a depresszióval nem diagnosztizáltak között számos nyelvi marker tekintetében. A depresszióval küzdők nagyobb arányban használnak **múlt idejű igéket** (ez jól illeszkedik az ún. *rumináció* lélektani sajátosságaihoz, amikor is már megtörtént, és így megváltoztathatatlan múltbeli eseményeken rágódunk); fekete-fehér vagy túláltalánosító gondolkodást jelző **általános névmásokat** (pl. mindig, minden, mindenki), amelyek egyben a kínzó jelen reménytelenségét és látszólagos megmásíthatatlanságát is hűen leképezik; **tagadó formulákat** (pl. soha, senki, nem) a negativisztikus önkép, világkép, jövőkép (más szóval a *depresszív kognitív triász*) megnyilvánulásaként; **E/I személyes névmásokat** (mely a depresszióra olyannyira jellemző fokozott énfókusz egyértelmű lenyomata a szövegben); tartalmilag pedig jelentősen több **szomorúságot, haragot** és **magányosságot** kifejező szót. Érdekes az is, hogy a depresszióval diagnosztizált személyek

bejegyzései terjedelmileg hosszabbak. Ez elsősorban a befelé fordulás, a belső élményvilágba zárkózás tendenciájával hozható összefüggésbe: a mindennapi személyes interakciók során bennük rekedt kommunikációs igény beszéd helyett írásban követel helyet magának.

A pszicholingvisztikai elemzés egy adott mentális zavar automatizált felismerésén túl kockázatbecslésre és akár megelőzésre is használható. Remek példa erre az az empirikus tapasztalat, hogy pusztán a közösségi média bejegyzések alapján megjósolható a szuicid kockázat (pl. Coppersmith és mtsai., 2018). A közeljövőben várhatóan számos közösségi oldal kínál majd felhasználói számára nyelvhasználaton alapuló profilozást, és az esetleges érintettség esetén súlyosságbecslést.

A saját logika börtönében?

Láthatjuk, hogy a nyelv még akkor is képes egy-egy pszichológiai problémát finoman jelezni, ha a nyelvhasználat mögött rejlő kognitív struktúra nem mutat szélsőséges patológiát. De mi történik, ha a zavar egyik legalapvetőbb ismérve éppen a kognitív folyamatok fellazulása, a valóságtól elszakadás, az elsodródás az emberi kapcsolatok szabályszerűségeitől?

Milyen nyelvi jelei vannak a pszichózisnak?

Mielőtt erre választ keresnénk, érdemes megvizsgálnunk, hogy a pszichotikus zavarok esetén milyen szinteken módosulhat a nyelvhasználat. Vegyük alapul a szkizofrénia nyelvi modelljét! A kommunikáció, és összességében a beszéd három fő pillére, a *tartalom* (1), a *produkción* (2) és a *percepción* (3) jól kivehetően torzul szkizofréniával élő betegeknél. 1) A beszéd tartalmát téveszmék, sokszor paranoid színezetű (pl. üldöztetéssel vagy lehallgatással kapcsolatos) gondolatok szövik át, melyekben sokszor kifejezetten fenyegető vagy erőszakos tematika rajzolódik ki (ez a szöveg szintjén például az agressziót kifejező szavak magasabb arányában érhető tetten). A téveszmék ugyanakkor korántsem függetlenek a szociokulturális közegtől. Több vizsgálat is alátámasztotta például, hogy a vallásos téveszmék gyakrabban fordulnak elő keresztény kultúrkörben, mint hindu, buddhista vagy muszlim közegben, vagy hogy a mágikus gondolkodáshoz köthető téveszmék a kistelepülések, a falvak lakói körében jóval meghatározottabbak, mint a nagyvárosokban élőkénél. 2) A beszéd produkcióját sajátos belső logika "szervezi" inkoherenssé. Mondhatnánk persze, hogy ezek a szövegek csak a külső megfigyelő számára inkoherensek, hiszen a szöveg előállítója valójában a saját szubjektív gondolatvilágában megadott szerkezet szerint építkezik, de ez nem fedné teljesen az igazságot. A pszichotikus zavarral küzdők ugyanis többnyire csapongó, gyakran változó elképzelések után kapkodnak, kétségbeesve próbálják elcsípni egy-egy ilyen gondolatfonál végét, ezáltal mintegy megelőznék a gondolatszálak teljes egymásba gabalyodását, a káosz elhatalmasodását. Hasonló elv mentén értelmezhetjük a pszichotikus betegek által kreált új szavak, a neologizmák törvényeit. Az 1970-es években két klinikus, LeVine és Conrad (1979) tett kísérletet, hogy szkizofrén pácienseik segítségével egyfajta neologizma-szótárat hozanak létre, melyben minden új, a betegek által gyártott szóhoz hozzárendelték annak hétköznapi nyelvi megfelelőjét. Olyan szópárookra gondolhatunk itt (dőlt betűvel szedve a neologizmákat), mint *memogener*-memorize (vagyis memorizálni), *orbynate*-orbit (keringeni), *manlas*-man (férfi), *baebardistraed*-excited (izgatott), vagy *abbliscotage*-soup (leves). E néhány példa alapján is látszik, hogy a neologizmák hol az eredeti szó vagy szótó kiegészítéseként, a forrásalaktól lényegesen nem eltérve jelennek meg, hol teljességgel elválnak az eredeti alaktól, és valami gyökeresen új szóalakot formálnak. Ebben, tapasztalataink alapján, jellemzően nincs következetesség, szabályszerűség, a neologizmák sokkal inkább ad hoc jelleggel teremődnek. És, akár csak a téveszmék szociokulturális meghatározottsága, a nyelvi kohézió szemantikai és grammatikai elemei esetén is komoly individuális különbségek adódhatnak szkizofrén betegek között. Erre talán a legékeesebb bizonyíték, hogy többnyelvű páciensek körében általában csak az egyik használt nyelv mutat

jelentős inkoherenciát (Hughes, 1981). 3) Végül, a beszéd percepciójának szintjén, szkizofréniával küzdők esetén gyakorta figyelhetőek meg „saját gyártású” észleleti anyagok, melyek hanghallások (és nem ritkán párbeszédet generáló hanghallások) formájában hoznak létre feldolgozandó szövegeket. Olyan, csak a beteg számára hallható hangok ezek, melyek az ún. auditoros/hallásos hallucinációk közé sorolhatók. Ezen hangok egy része zörej jellegű, de az esetek többségében emberi beszéd formájában jelentkeznek. És ha már az imént szóba kerültek a több nyelvet beszélők: Narcyz Lukianowicz, a bristoli Barrow Kórház ukrán származású pszichiátere arra az izgalmas megállapításra jutott szkizofrén páciensekkel folytatott több évtizednyi munkája során (1962), hogy a hanghallásokról beszámoló poliglott betegek általában a második (vagy többedik tanult) nyelvükön hallják a fenyegető, rémisztő hangokat, míg az őket óvó, nyugtató hangok szinte mindig az anyanyelvükön szólnak hozzájuk.

Pszichózis és szövegbányászat

Milyen nyelvhasználati jegyek jellemzik a tartós pszichotikus zavarban szenvedők táborát? A gépi tanulás módszerét alkalmazó tanulmányok (pl. Bedi és mtsai, 2015; Corcoran és mtsai., 2018; Mota és mtsai., 2018; Rezaii és mtsai., 2019) elsősorban a szemantikus koherencia és néhány szintaktikai jegy alapján több mint 90%-os megbízhatósággal tudták beazonosítani a fennálló pszichózist, vagy akár előre jelezni a későbbi pszichotikus epizódokat a zavar szempontjából rizikócsoportha (értsd: genetikai, családi halmozódás) sorolt serdülők körében. A szemantikus koherencia pszichotikus betegek esetén jól kivehetően sérül. Az általuk létrehozott szöveg (legyen az írott vagy beszélt nyelv) alacsony szemantikai sűrűséget mutat, vagyis egészséges kontrollszemélyekhez viszonyítva jóval kisebb az egy mondaton belül megjelenő értelmes vagy egymásból logikusan következő szavak és kifejezések aránya, illetve az egyes szavak kevésbé rendezhetőek egységes, koherens szemantikai blokkokba. Izgalmas megfigyelés továbbá, hogy a felnőtt pszichotikus páciensek iskolai végzettségüktől függetlenül általában a gyermekekéhez hasonló szövegeket hoznak létre, már ami az egy adott témán belül megjelenő lexikális változatosságot és szemantikai komplexitást illeti. Mintha gondolkodásuk tartós zavara megakadályozná a szövegszerkesztés életkorral vagy iskolázottsággal összefüggő természetes fejlődését. A szintaktikai jegyek közül kiemelhető, hogy az organikus eredetű pszichotikus zavarral küzdők általában rövidebb kifejezéseket (alacsonyabb karakterszám), kevesebb határozószót (pl. az, azokat, amit) és birtokos névmást használnak, illetve gyakrabban hagyják félbe mondataikat.

Hogyan tudnánk ezeket az eredményeket a lélektan nyelvére fordítani? Érdekes abból kiindulnunk, hogy a pszichózis, és kiváltképpen a szkizofrénia, a szubjektív élményvilágba zárkózás betegsége, ezzel párhuzamosan pedig az emberi kapcsolatoktól távolodás, az elmagányosodás kórképe is. Számos – kevésbé szélsőségesen medikalizált gondolkodást mutató – országban (pl. Skandinávia) nagyobb teret kap a szkizofrénia közösségi terápiája, a betegek reszocializációja, társas kompetenciájának fejlesztése. Ha érzelmileg és a gondolatok síkján is távolodom társas (és akár tárgyi) környezetemtől, az törvényszerűen kihat arra is, hogy nyelvhasználatommal mit emelek ki a külső, objektív valóságból. A birtokos névmások és a rámutatás szimbolikus funkciójával rendelkező határozószók emberi tulajdon- és kapcsolatrendszer leírását szolgálják, általuk interakciót is létesíthetünk másokkal, hiszen a rámutatás már a csecsemők kommunikációs kísérlete. Ezeknek a szófajoknak az alacsony megjelenési aránya tehát értelmezhető a világból kivonulás tendenciájával is. Hasonlóan, a szemantikai töredezettség is elszakadás a nyelvi szabályszerűségektől, a megszokottól eltérő szövegalkotási logika manifesztuma. És ebbe a tárgykörbe sorolható az is, hogy a szkizofrén betegek kommunikációjában, különösen érzelmileg megterhelő, negatív témák esetén, megszorodik az ún. *referencia-hiba* valószínűsége, azaz kevésbé képesek egyértelműen, a

másik fél számára is érthetően átadni egy-egy kommunikációs üzenetet, mert utalásaik zavarosak, előzmény nélküliek vagy összefüggéstelenek maradnak (Rubino és mtsai., 2011).

Nyelv-diagnosztika?

De megállja-e helyét a pszicholingvisztika a pszichiátria diagnózisközpontú, tünetcsoportokra fókuszáló paradigmarendszerében? A világ vezető pszichiátriai szaklapjában, a *World Psychiatry*-ban közölték a tanulmányt (Corcoran és mtsai., 2018), melyben a szerzők bizonyították, hogy a lingvisztikai sajátosságok megbízhatóbban jelzik a pszichotikus zavarok későbbi súlyosbodását, mint a hagyományos klinikai értékelőskálák. A klinikumban dolgozók így olyan eszközt kapnak, amely nagyfokú érzékenységgel szolgálja mindennapi diagnosztikai ténykedésüket, ráadásul – minthogy nyelvhasználatunk árnyalt jellemzőit kevésbé kontrolláljuk – csökken a szimulációs/disszimulációs valószínűség is. A nyelv a projektív tesztekhez hasonlóan képes megragadni olyan mélylélektani elakadásokat is, amiket a papír-ceruza tesztek sokszor képtelenek azonosítani. A pszicholingvisztika abban is segítségünkre lehet, hogy hatékonyabban tudjunk elválasztani tünettanukban átfedést mutató kórállapotokat (más szóval: a differenciáldiagnosztikai felhasználás is elképzelhető). Sok esetben nehézséget jelent egyetlen találkozás alkalmával különbséget tenni a drog indukálta pszichózis és a tartós, organikus eredetű pszichotikus zavarok között, ám ha tudjuk, hogy a krónikusan fennálló pszichotikus állapotok esetén nagyobb arányban torzul a mögöttes kognitív működés, számíthatunk rá, hogy mindez a nyelvhasználat formájában is tetten érhető.

Akárhogyan is, nyelvünk ezerfényű kincsesbánya, lélektani folyamataink finoman hangolt csatornája, így éppen itt az ideje, hogy a mainstream pszichiátria is felfigyeljen a pszicholingvisztikában rejlő lehetőségekre. Ahogyan Tótfalusi István fogalmazott, *„beszédünkben éppen az a jó, hogy pontos képet, láttelepet ad a beszélő lelkületéről”*.

Irodalom

- Bedi, G., Carrillo, F., Cecchi, G., Slezak, D.S., Sigman, M., Mota, N.B., Ribeiro, S., Javitt, D.C., Copelli, M., Corcoran, C.M. (2015). Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophrenia*, 1, 15030. <https://doi.org/10.1038/npjSchz.2015.30>
- Cacheda, F., Fernandez, D., Novoa, F. J., & Carneiro, V. (2019). Early Detection of Depression: Social Network Analysis and Random Forest Techniques. *Journal of Medical Internet Research*, 21(6), e12554. <https://doi.org/10.2196/12554>
- Coppersmith, G., Leary, R., Crutchley, P., & Fine, A. (2018). Natural Language Processing of Social Media as Screening for Suicide Risk. *Biomedical Informatics Insights*, 10, 1178222618792860. <https://doi.org/10.1177/1178222618792860>
- Corcoran, C. M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., Javitt, D. C., Bearden, C. E., & Cecchi, G. A. (2018). Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry*, 17(1), 67–75. <https://doi.org/10.1002/wps.20491>
- Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preotiuc-Pietro, D., Asch, D. A., Schwartz, H. A. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences of the United States of America*, 115(44), 11203–11208. <https://doi.org/10.1073/pnas.1802331115>
- Hughes, G.W. (1981). Neuropsychiatric aspects of bilingualism: a brief review. *British Journal of Psychiatry*, 139, 25-28. doi: 10.1192/bjp.139.1.25
- LeVine, W.R., Conrad, R.L. (2017). The classification of schizophrenic neologisms. *Psychiatry*, 42(2), 177-181. doi: 10.1080/00332747.1979.11024022.

- Lukianowicz, N. (1962). Auditory hallucinations in polyglott subjects. *Psychiatria et Neurologia (Basel)*, 143, 274–294.
- Mota, N.B., Sigman, M., Cecchi, G., Copelli, M., Ribeiro, S. (2018). The maturation of speech structure in psychosis is resistant to formal education. *npj Schizophrenia*, 4, 25. <https://doi.org/10.1038/s41537-018-0067-3>
- Rezaii, N., Walker, E., & Wolff, P. (2019). A machine learning approach to predicting psychosis using semantic density and latent content analysis. *NPJ Schizophrenia*, 5(1), 9. <https://doi.org/10.1038/s41537-019-0077-9>
- Rubino, I. A., D'Agostino, L., Sarchiola, L., Romeo, D., Siracusano, A., & Docherty, N. M. (2011). Referential failures and affective reactivity of language in schizophrenia and unipolar depression. *Schizophrenia Bulletin*, 37(3), 554–560. <https://doi.org/10.1093/schbul/sbp108>
- Trifan, A., Antunes, R., Matos, S., Oliveira, J. L. (2020). Understanding Depression from Psycholinguistic Patterns in Social Media Texts. *Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14–17, 2020, Proceedings, Part II*, 12036, 402–409. https://doi.org/10.1007/978-3-030-45442-5_50
- We Are Social, Hootsuite (2021). Digital 2021-Global Overview Report. Online elérés: <https://wearesocial.com/blog/2021/01/digital-2021-the-latest-insights-into-the-state-of-digital> (elérés dátuma: 2021. május 14.)