

Новак В. Р.,

Східноєвропейський національний університет ім. Лесі Українки, м. Луцьк

ДО ПИТАННЯ ВИКОРИСТАННЯ НЕВЕРБАЛЬНИХ КОМПОНЕНТІВ В КОМУНІКАЦІЇ «ЛЮДИНА-КОМП'ЮТЕР»

У статті висвітлено сучасні розвідки в галузі людино-машинної взаємодії, зокрема функціонування мультимодальних інтерфейсів користувача. Розглянуто кінетичний компонент невербальної комунікації як центральний в сфері створення нового покоління інтерфейсів. Наведено класифікації жестів, що є найбільш релевантними при взаємодії людини і комп'ютера, та запропоновано приклади їх застосування, а також перспективи подальших розвідок у даній галузі.

Ключові слова: взаємодія «людина-комп'ютер», мультимодальний інтерфейс, кінетичний невербальний компонент, ергосеміотика.

В статье освещаются современные исследования в области человеко-машинного взаимодействия, в частности функционирование мультимодальных пользовательских интерфейсов. Рассматривается кинетический компонент невербальной коммуникации как центральный в сфере создания нового поколения интерфейсов. Приводятся классификации наиболее релевантных при взаимодействии человека и компьютера жестов, предлагаются примеры их использования, а также перспективы дальнейших исследований в данной области.

Ключевые слова: взаимодействие «человек-компьютер», мультимодальный интерфейс, кинетический невербальный компонент, эргосемиотика.

Up-to-date investigations in the sphere of human-computer interaction are highlighted in the article, in particular functioning of multimodal user interfaces. Kinesic component of nonverbal communication is studied, being a leading one in creation of the new generation of interfaces. The most relevant to human-computer interaction classifications of gestures are provided, the examples of their usage are given, as well as the prospects of further research in this field.

Key words: human-computer interaction, multimodal interface, kinesic nonverbal component, ergosemiotics.

Технічні досягнення ХХ століття в області зв'язку й інформації призвели до того, що у світі склалося особливе віртуальне середовище взаємодії людей за допомогою комп'ютера. У зв'язку із цим актуальними стають проблеми ефективної комунікації між людиною та комп'ютером які на даному етапі намагаються вирішити як комп'ютерні інженери, програмісти, так і лінгвісти всього світу. Так, авторами досліджень взаємодії «людина-комп'ютер» є Д. Ятес, В. Орликовський, Д. Гумперт, Т. Нільсен, Л. Спроул, С. Кейслер, С. Хилтз, М. Туроф, Д. Норман та ін.. Питання проблем взаємодії вербальних та невербальних чинників в такій комунікації затркують в своїх роботах Р. Райс, Д. Лав, Л. Спроул, Д. МакНіл, Д. Росіні, А. Малдер, А. Кендон та ін.. Зараз, в період загальної комп'ютеризації та глобальної інформатизації, на перше місце виходить проблема створення такого інтерфейсу користувача, який не тільки б задовольняв усі вимоги до ефективної взаємодії людини з комп'ютером, а й значно б її полегшував та покращував результативність завдяки якнайбільшій відповідності програмного забезпечення природній мові людини (у її як вербальному, так і невербальному вираженні) та його відповідності всім особливим потребам користувача.

Відомий спеціаліст в галузі комп'ютерної семіотики Д. Норман вважає, що для створення якісного інтерфейсу користувача комп'ютера необхідно володіти: 1) мовами програмування та інформаційних технологій; 2) знанням про людей, особливості їхнього спілкування між собою та з комп'ютерною системою; 3) знанням тієї прикладної галузі, задля вирішення завдань якої створюється програмний продукт [1, с. 149]. При вивченні інтерфейсу користувача слід брати до уваги природу взаємодії людини з комп'ютерною системою, а також особу самого користувача: його знання, нахили, здібності, потреби, тощо.

З 1988 року в науковій літературі починають обговорюватися основні положення нової науково-практичної спеціалізації, що займається дослідженням проблем, спільних для семіотики, лінгвістики та ергономіки, – ергономічної семіотики (надалі ергосеміотики). Вона спрямована на розробку принципів проектування та створення ефективних та прийнятних мов людино-машинної, зокрема людино-комп'ютерної взаємодії. В основу ергосеміотики покладено ідеї Ч. С. Пірса про те, що розуміння природи іконічних та індексальних знаків та символів може сприяти належному використанню властивих людині мов у створенні штучних знакових систем. Більшість властивих людині кодів з'явилась ще у довербальному періоді еволюції засобів спілкування. До них можна віднести такі знакові системи як мова рухів, дій, жестів, образів, емоційних станів, іконічна мова, музичний та кольоро-коди тощо. У зв'язку з тим, що ці мови та коди не є добре вивченими, до завдань ергосеміотики належать дослідження цих семіотичних утворень у двох напрямках: а) проектування складних семіотичних утворень; б) сполученості якостей, функцій та особливостей штучних мов людино-комп'ютерної взаємодії з характеристиками нових засобів та приладів для введення та виводу інформації.

Останнім часом активно розвиваються засоби спілкування у віртуальній реальності, Інтернеті тощо. На думку Г. Я. Узилевського, цей період відрізнятиметься безпосередньою взаємодією користувачів із комп'ютерами, коли мови та коди, що з'явилися у довербальному періоді, усний варіант природної мови та нові засоби введення та виводу інформації будуть відігравати домінуючу роль; з'являтимуться нові знакові засоби, побудовані на симбіозі цих мов та кодів [4, с. 5]. В контексті ергосеміотики конкретний інтерфейс користувача визначається як інтерактивна багаторівнева інформаційна система, утворена на основі вже наявних знакових систем (природної мови, іконічної мови, кольорових та музичних кодів тощо), а також штучних мов, що еволюційно розвиваються, та гнучких способів створення програмних і технічних засобів. Для вдосконалення комунікації «людина-

комп'ютер» в її основу має бути закладена не лише модель обробки інформації в процесі спілкування людина-людина, а й модель породження рухів та дій як елементи вербального та невербального мислення, тісно пов'язані між собою прямими та зворотними зв'язками, багатоярусністю та системною організацією. Така модель сприяє введенню в людину-комп'ютерну взаємодію нових метафор, проектуванню сучасних комп'ютерних систем та дослідженню ментального світу користувача.

Характерною рисою розвитку сучасних комп'ютерних технологій є активне впровадження мульти-медіа технологій в інформаційні системи як для загальних завдань представлення інформації, так і для забезпечення взаємодії користувача з програмним середовищем. В умовах, коли комп'ютерні технології стають доступними дедалі ширшому колу людей, найбільш перспективним стає напрямок комп'ютерного діалогу з елементами інтелектуального інтерфейсу, який використовує як природну мову, так і невербальні засоби комунікації. Застосування невербальних мовних засобів в спеціалізованих програмних середовищах дозволяє підвищити ефективність їх використання. Вдосконалений таким чином комп'ютерний діалог розширює сигналізаційні можливості інтерфейсу як кількісно, так і якісно. Інтеграція невербальних компонент комунікації у людину-машинний інтерфейс сприяє підвищенню швидкості та повноти засвоєння інформації користувачами програм, що є особливо важливим для таких класів програмного забезпечення, як системи контролю даних, комунікаційні програми, операційні системи, навчальні, розвиваючі та довідкові системи.

З 90-х років минулого століття розпочалися дослідження нового покоління інтерфейсів користувача, які не використовують меню, форми та панелі інструментів; замість них для визначення специфікацій операцій робиться наголос на навчальні приклади, жести та розпізнавання мовлення. Такі інтерфейси ще не набули широкої популярності, проте спеціалісти різних прикладних галузей активно працюють над їх вдосконаленням. Виходячи з цього, в нашому дослідженні ставимо за мету висвітлення прикладних аспектів кінесики – однієї із центральних галузей невербальної семіотики, бо саме вона зараз проходить процес найбільшої інтеграції у людину-машинну комунікацію.

Для вирішення глобальної проблеми людину-машинної взаємодії в інтерфейсах користувача дедалі більше використовуються додаткові види каналів передачі інформації (мовлення, артикуляція губ, жести, напрямок погляду тощо). Такий спосіб взаємодії отримав назву «мультимодальна взаємодія»; вона реалізується шляхом мультимодальних інтерфейсів. Саме такі інтерфейси є властивими міжлюдському спілкуванню. В таких програмах користувач обирає, який канал, для передачі якого типу інформації йому найбільш зручно використовувати у даний момент. Такі інтерфейси дозволяють забезпечити найбільш ефективну та природну для людини взаємодію з різноманітними автоматизованими засобами керування та комунікації. В мультимодальних системах інформація від аудіо, відео, тактильних та інших комунікативних каналів безперервно обробляється, утворюючи реальне чи віртуальне оточення, що дозволяє задовольнити бажання користувача та оперативно адаптуватися до контексту [2, с. 301]. Отже, моделювання мультимодальної поведінки людини є основою для проектування інтелектуальних систем природної взаємодії. Мультимодальні мовленнєві системи (аудіовізуальні) є спробою досягнути простоти комунікації «людина-комп'ютер», поєднуючи автоматичне розпізнавання мовлення з іншими невербальними засобами, а також інтегруючи невербальні засоби із синтезом мовлення задля покращення методу введення інформації в мультимодальному пристрої. Таким чином, застосування мультимодальності дозволяє розширити спектр користувачів системи та потенційно забезпечує адаптивність до специфічних умов функціонування.

Можливість за допомогою мультимодальних систем обробляти вхідні модальності паралельно сприяє покращенню ефективності взаємодії людини та комп'ютера, особливо при маніпулюванні графічною інформацією. Оскільки користувачі сприймають інформацію на екранах комп'ютерів візуально (та частково на слух), і використовують рухи для проведення найпростіших операцій у графічному інтерфейсі комп'ютера (такі як клік, подвійний клік, пересування, що супроводжуються виконанням певних команд), тому можна стверджувати, що оптико-кінетичні засоби комунікації превалюють у спілкуванні людини з комп'ютером. Говорячи про рухи у спілкуванні з комп'ютером, варто розглянути природні жести людини в рамках функцій, які вони виконують. На сьогоднішній день існує багато різних класифікацій жестів, запропонованих такими дослідниками невербальної комунікації як С. Cadoz, D. McNeill, D. Efron, P. Ekman, B. Rimé і L. Schiaratura, A. Kendon та ін.; результати їх досліджень висвітлив у своїй роботі А. Mulder [6]. Стосовно застосування жестів у комунікації «людина-комп'ютер», на нашу думку, заслуговує на увагу класифікація, запропонована Д. МакНілом. Він виділяє жести:

- іконічні (що відображають певний об'єкт чи подію та тісно співвідносяться з семантичним змістом мовлення),
- метафоричні (як іконічні, але відображають абстрактні поняття),
- дейктичні (що вказують на когось чи щось конкретне чи то абстрактне) – це жести-вказівки, що застосовуються для привертання уваги,
- жести-акценти чи жести-наголоси (що мають лише дві фази (вгору/вниз, всередину/назовні), наголошуючи на значущості слова чи фрази, яких вони стосуються).

Говорячи про взаємодію людини з комп'ютером, в першу чергу згадують саме про дейктичні жести, що не дивно через те, що ці жести використовуються фактично у всіх доменах людину-машинної взаємодії (віртуальна реальність, мобільні телефони, планшети тощо). В роботі R. Bolt «Put that there» (1980) дейктичні жести було використано разом із командами вводу голосом для вказівки на розташування предметів на великому екрані. Ця робота є одним із перших досліджень з мультимодальних систем, в якому демонструвалася концепція використання мовлення та ручного маніпулювання об'єктом. Починаючи з цієї роботи, почали з'являтися бібліотеки та алгоритми для підтримки нових видів взаємодії, а також реальні програми (додатки). З базової концепції R. Bolt розвинулася загальна ідея мультимодальних систем, а емпірична робота з їх створення сприяла розробці теорії HSMI (human-computer multimodal interaction), яка наразі забезпечує основу для написання посібників з проектування таких систем.

Варто також розглянути класифікацію жестів, запропоновану F. Quek, в якій виділяється п'ять основних типів жестів у взаємодії «людина-комп'ютер»: жестикуляційні, маніпуляційні, сигналізаційні, дейктичні та мовоподібні.

Зупинимось на різноманітних маніпуляційних жестах, які існують у взаємодії людини з комп'ютером. За визначенням F. Quek, призначенням маніпуляційного жесту є керування певним об'єктом із використанням зв'язку між фактичним рухом рук та маніпуляцією об'єктом [7]. Маніпуляція відбувається як із використанням певних пристроїв для маніпуляції (наприклад, комп'ютерна мишка, стилус), так і з використанням вільних рухів рук для імітації певних дій із фізичним об'єктом; як в інтерфейсах віртуальної реальності, так і в «фізичних» інтерфейсах. До маніпуляційних жестів відносяться жести-вказівки, двовимірні та тривимірні жести. Для вказування використовуються спеціальні указки, палець руки чи світлове перо та активний (сенсорний) екран. Двовимірні жести є рухами на площині, коли маніпулюють безпосередньо курсором, вікном чи іншим об'єктом на екрані. Традиційно для маніпулювання у графічних інтерфейсах користувача використовується мишка, стилус, джойстик, трекбол чи інші засоби. Тут слід зауважити, що такі дії як перетягування та клік не вважаються жестами, доки користувач не виконає певний жест, який інтерпретуватиметься як команда. Тривимірні жести є результатом руху пальців, рук чи голови у тривимірному просторі.

Для створення систем інтерактивної взаємодії «людина-комп'ютер» важливі також класифікації, в яких враховано зв'язок між мовленням та жестами. Так, A. Kendon запропонував розташувати жести в порядку залежності від зв'язків мовлення/жест таким чином:

- жестикуляція (жести наголосу, жести-зв'язки);
- мовоподібні жести (іконічні/образотворчі жести);
- міміка, пантоміма (пантомімічні жести);
- дейктичні жести (вказівні);
- мова жестів (мова глухонімих).

Чим ближче до останньої групи (мова жестів), тим менше необхідність у супроводжувальному мовленні. Жести стають більше подібними до мови. Так, наприклад, мова глухонімих враховує більшість семантичних та синтаксичних особливостей мовлення, тому не потребує самого мовлення для її інтерпретації. Проте іконічні жести (образотворчі) важко зрозуміти без додаткової інформації.

Зупинимось докладніше на жестикуляції та міміці, представлених у класифікації. У взаємодії людини з комп'ютером жестикуляція, як найбільш природна форма жестів, застосовується у мультимодальних інтерфейсах, де мовлення та жести розглядаються як єдиний засіб комунікації. Оскільки всі люди при мовленні так чи інакше застосовують жестикуляцію, актуальним постає її врахування при створенні анімаційних героїв чи роботів, за посередництвом яких відбувається комунікація.

На думку В. Є. Тимофєєва, моделювання на комп'ютері саме мімічного апарату людини є найбільш візуально ефективним засобом удосконалення людино-машинного інтерфейсу. Мова міміки є досить складною та здатна передавати великі обсяги змістовно насиченої та різноманітної інформації. Тому одним із завдань лінгвістів, які займаються вивченням цього питання, є надання розробникам програмного продукту наукової бази для створення мімічного інтерфейсу, що може стати початком формування комп'ютерної моделі співрозмовника, яка інтегруватиметься у людино-машинний інтерфейс та включатиме як міміку, так і інші невербальні мовні засоби, а також мовлення [3]. Необхідно зауважити, що мімічний компонент у комунікації, напевно, вперше відобразився у створенні та широкому застосуванні «смайликів», що моделюють людські емоції, без яких деякі користувачі вже не можуть повноцінно сприймати певну інформацію. Тобто, моделювання мімічних систем почалося саме із створення графічних символів для позначення людських емоцій, які надалі систематизуються у бази даних, що використовуються для візуалізації міміки. Сучасні ж комп'ютерні програми вже пропонують нам спілкування не просто із зображенням людини на екрані, а з її «оживленим» варіантом, який рухає губами, бровами, головою, має певний вираз обличчя та здатен виражати різні емоції. Одним із найбільш перспективних напрямків досліджень у цій галузі стане трансформування міміки довільного людського обличчя, що вводиться в систему як графічний об'єкт безпосередньо в процесі її роботи. Актуальним є також застосування мімічної системи у навчальних програмах, а саме для оцінювання успішності, тобто такого моделювання функцій вчителя, без якого важко розраховувати на ефективність навчання.

Іноземні дослідники (W. Gao, C. Maggioni, N. Hataoka, K. Bohm та ін.), які вивчають мультимодальну взаємодію, що включає не лише жести та мовлення, але й вирази обличчя (міміку), створили систему, спроможну ідентифікувати кожну модальність окремо. З точки зору дослідників людської поведінки, недоліки такої системи полягають у тому, що при перевірці та зіставленні інформації, отриманої окремо з кожної модальності, виявляється, що такі види інформації дублюють одна одну. Доцільним є створення системи, в якій окремі модальності, виражаючи специфічні аспекти комунікації, не лише б підтверджували правильність інтерпретації інших, а й доповнювали б одна одну.

Розглядаючи кінетичний компонент невербальної комунікації людини з комп'ютером, важливо відзначити той факт, що людина має дві руки. Оскільки 90% людей є праворукими, то й більшість інтерфейсів користувача розробляються з урахуванням того, що майже всі дії виконуватимуться самою правою рукою. Проте дослідники у цій галузі (J. Napier, D. Kimura, W. Buxton та ін.) наголошують, що хоча ліва півкуля головного мозку людини (що відповідає за діяльність правою рукою) є важливою при свідомому виборі рухів, пальці лівої руки краще виконують ізольовані рухи, ніж пальці правої. У. Бакстон та Б. Майерс, в свою чергу, експериментально довели, що залучення обох рук при редагуванні тексту та одночасній навігації по документу сприяє значному покращенню продуктивності виконання завдання та зменшенню витраченого часу, при чому не лише у досвідчених користувачів, а й у новачків. Таким чином учені доводять, що необхідні навички моторики є природними для людини незалежно від міри її обізнаності з комп'ютерними інтерфейсами і з легкістю застосовуються нею в комунікації «людина-комп'ютер» [5, с. 409-412].

Кінетичний компонент комунікації також відіграє важливу роль у діалозі з комп'ютером користувачів із вадами зору, для яких вже було винайдено брайлевський дисплей та брайлевську клавіатуру. Проте в цьому випадку ефективнішою стає комбінація аудіальних та тактильних чинників. Для користувачів із вадами слуху нещодавно було сконструйовано спеціальну сенсорну рукавичку, яка здатна перекладати жестову мову у текст на дисплеї.

В ході досліджень було виявлено, що приблизно 95% користувачів надають перевагу мультимодальній взаємодії в роботі, наприклад, із інтерактивними картами, вільно використовуючи як мовленнєве введення, так і вказівки пальцем чи світлове перо. Застосування мультимодальних інтерфейсів сприяє збільшенню швидкості виконання візуально-просторових завдань на 10% [2, с. 311-312].

Найбільші функціональні можливості мають системи, які використовують мультимодальне подання інформації, тобто застосовують, паралельно або послідовно, кілька різних комунікаційних каналів (слуховий, зоровий, тактильний та ін.). Завдяки цьому мультимодальні системи дають змогу надзвичайно збільшити обсяг передавання інформації, що є головною їх перевагою. Мультимодальність спілкування підвищує ефективність навчання завдяки притаманному людині механізмові паралельного опрацювання інформації, що надходить різними каналами. Потрібно проектувати систему таким чином, щоб всі сенсорні модальності доповнювали, а не дублювали і заважали одна одній, примушуючи переключати увагу і відволікати внутрішні ресурси людини від переробки інформації. Доцільно вважати мультимодальність необхідною умовою психологічної природності та комфортності спілкування.

Розробка системних принципів створення комфортного та ефективного діалогу людини з комп'ютером передбачає вивчення основних компонентів системи «людина-комп'ютер», а саме психологічних, семіотичних, ергономічних та інших аспектів процесу комунікації. В статті ми проаналізували ефективність залучення та використання у сучасних комп'ютерних програмах жестів та міміки задля більш ефективної взаємодії людини з комп'ютером. Незважаючи на те, що окремі дослідники виводять на перший план різні невербальні чинники, аналіз їх застосування у людино-машинному спілкуванні показав, що важливість того чи іншого компонента для користувача є, по-перше, ситуативно-зумовленою, а по-друге, залежить як від особливих потреб користувача (наприклад, певних фізичних вад), так і від кінцевих цілей комунікації.

Зараз рішенням проблем ефективної взаємодії людини з комп'ютером у природній формі займаються спеціалісти різних наукових галузей. Інженери, математики, програмісти розробляють більш ергономічні програмні та апаратні засоби керування комп'ютерами. В свою чергу лінгвісти, фізіологи, психологи вивчають поведінку людини, а також механізми сприйняття, обробки інформації задля того, щоб зрозуміти специфіку процесів мислення. Отже, на нашу думку, на даний час виглядає найбільш перспективним міждисциплінарний підхід до вирішення проблем побудови ефективних та природних інтерфейсів.

Література:

1. Агеев В. Н. Семиотика. – М. : Весь Мир, 2002. – 256 с.
2. Ронжин А. Л., Карпов А. А. Многомодальные интерфейсы: основные принципы и когнитивные аспекты / А. Л. Ронжин, А. А. Карпов // Труды СПИИРАН. Вып. 3, т. 1. – СПб. : Наука, 2006. – С. 300-319.
3. Тимофеев Е. В. Моделирование мимической компоненты обучающих информационных систем, 2001. – [Електронний ресурс] – Режим доступу : <http://www.disscat.com/content/modelirovanie-mimicheskoi-komponenty-obuchayushchikh-informatsionnykh-sistem>
4. Узилевский Г. Я. Основания, цели, задачи и потенциал эргономической семиотики. – [Електронний ресурс] – Режим доступу : <http://stra.teg.ru/library/strategics/7/10/0>
5. Buxton W. The Natural Language of Interaction: A Perspective on Non-Verbal Dialogues / B. Laurel (Ed.) // The Art of Human-Computer Interface Design, Reading, – MA : Addison-Wesley, 1990. – С. 405-416.
6. Mulder A. Hand Gestures for HCI, 1996. – [Електронний ресурс] – Режим доступу : <http://xspasm.com/x/sfu/vmi/HCI-gestures.htm#classify>
7. Quek F., McNeill D., Bryll B., Duncan S., Ma X., Kirbas C., McCullough K., Ansari R. Multimodal Human Discourse: Gesture and Speech, // ACM Transactions on Computer-Human Interaction, vol. 9, no. 3, 2002. – P. 171-193.